

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta elektrotechniky  
a komunikačních technologií

DIPLOMOVÁ PRÁCE



**VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ**

BRNO UNIVERSITY OF TECHNOLOGY

**FAKULTA ELEKTROTECHNIKY**

**A KOMUNIKAČNÍCH TECHNOLOGIÍ**

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

**ÚSTAV TELEKOMUNIKACÍ**

DEPARTMENT OF TELECOMMUNICATIONS

**MIKROFONNÍ POLE MALÝCH ROZMĚRŮ PRO ODHAD  
SMĚRU PŘÍCHÁZEJÍCÍHO ZVUKU**

SMALL-SIZE MICROPHONE ARRAY FOR ESTIMATION OF DIRECTION OF ARRIVAL OF SOUND

**DIPLOMOVÁ PRÁCE**

MASTER'S THESIS

**AUTOR PRÁCE**

AUTHOR

**Bc. Ladislav Kubišta**

**VEDOUCÍ PRÁCE**

SUPERVISOR

**doc. Ing. Jiří Schimmel, Ph.D.**

**BRNO 2020**



# Diplomová práce

magisterský navazující studijní obor **Audio inženýrství**

Ústav telekomunikací

**Student:** Bc. Ladislav Kubišta

**ID:** 174460

**Ročník:** 2

**Akademický rok:** 2019/20

**NÁZEV TÉMATU:**

## Mikrofonní pole malých rozměrů pro odhad směru přicházejícího zvuku

### POKYNY PRO VYPRACOVÁNÍ:

Prostudujte metody odhadu směru přicházejícího zvuku pomocí mikrofonních polí malých rozměrů založené na měření fázového posunu, korelace a energie a tyto metody realizujte v prostředí Matlab. Navrhněte a zkonstruujte více mikrofonních polí malých rozměrů a proveďte srovnávací měření ve volném poli a v reálném prostředí pro vybrané typy zvukových signálů, např. výstřel, rozbití skla a siréna. Zaměřte se také na možnost a míru potlačení okolních zvuků a přesnost při velmi malých časových diferencích signálu v jednotlivých mikrofonech.

### DOPORUČENÁ LITERATURA:

- [1] Dan Li., Yu Hen Hu, "Energy-Based Collaborative Source Localization Using Acoustic Microsensor Array," EURASIP Journal on Applied Signal Processing, vol. 2003, no. 4, pp. 321-337, 2003.
- [2] Svaizer, P., Matassoni, M., Omologo, M. "Acoustic Source Location in a Three-Dimensional Space Using Crosspower Spectrum Phase," IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'97) - Volume 1, 1997.

**Termín zadání:** 3.2.2020

**Termín odevzdání:** 1.6.2020

**Vedoucí práce:** doc. Ing. Jiří Schimmel, Ph.D.

**prof. Ing. Jiří Mišurec, CSc.**  
předseda oborové rady

### UPOZORNĚNÍ:

Autor diplomové práce nesmí při vytváření diplomové práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č.40/2009 Sb.

## ABSTRAKT

Tato diplomová práce se zabývá detekcí směru přicházejícího zvuku pomocí analýzy zvuku v malých mikrofonních polích. Využívá metod založených na časovém zpoždění signálu, dále na energickém poklesu či rozdílné fázi signálu. Práce se zaměřuje na zjišťování úhlu pomocí malých časových diferencí signálu. V závěru jsou shrnuty výsledky měření určování směru zvuku nahraného laboratorních podmínkách a reálném prostředí. Veškeré výpočty byly provedeny pomocí simulačního prostředí Matlab.

## KLÍČOVÁ SLOVA

Direction of Arrival (DOA), Time Difference of Arrival (TDOA), Frequency of Arrival (FDOA), Cross-correlation, Generalized Cross-Correlation (GCC), Phase Transform (PHAT), Maximum Likelihood (MLM), B-format of signal, Short-Time Fourier Transform (STFT), Beamforming, small-size microphone array, Matlab, signal processing.

## ABSTRACT

This thesis describe detection of direction receiving sound with small-size microphone array. Work is based on analyzing methods of time delay estimation, energy decay or phase difference signal. Work focus mainly on finding of angle of arrival in small time difference. Results of measuring, as programming sound, so sound recorded in laboratory conditions and real enviroment, are mentioned in conclusion. All calculations were done by platform Matlab.

## KEYWORDS

Direction of Arrival (DOA), Time Difference of Arrival (TDOA), Frequency of Arrival (FDOA), Cross-correlation, Generalized Cross-Correlation (GCC), Phase Transform (PHAT), Maximum Likelihood (MLM), B-format of signal, Short-Time Fourier Transform (STFT), Beamforming, small-size microphone array, Matlab, signal processing.

KUBIŠTA, Ladislav *Mikrofonní pole malých rozměrů pro odhad směru přicházejícího zvuku*: diplomová práce. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací, 2020. 79 s. Vedoucí práce byl doc. Ing. Jiří Schimmel, Ph.D.



## PROHLÁŠENÍ

Prohlašuji, že svou diplomovou práci na téma „Mikrofonní pole malých rozměrů pro odhad směru přicházejícího zvuku“ jsem vypracoval(a) samostatně pod vedením vedoucího diplomové práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor(ka) uvedené diplomové práce dále prohlašuji, že v souvislosti s vytvořením této diplomové práce jsem neporušil(a) autorská práva třetích osob, zejména jsem nezasáhl(a) nedovoleným způsobem do cizích autorských práv osobnostních a/nebo majetkových a jsem si plně vědom(a) následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

Brno .....

.....  
podpis autora(-ky)

## PODĚKOVÁNÍ

Rád bych poděkoval vedoucímu diplomové práce panu Ing. Jiřímu Schimmelovi, Ph.D. za trpělivost, pomoc při řešení složitějších problémů, konzultace a podnětné návrhy k práci.

Brno .....

.....

podpis autora(-ky)

# OBSAH

Úvod	12
<b>1 Teoretická část práce</b>	<b>13</b>
1.1 Zvuk	13
1.1.1 Šíření zvuku prostorem	13
1.1.2 Akustický tlak	13
1.1.3 Akustická impedance	13
1.1.4 Typy zvukových polí	14
1.1.5 Typy mikrofonních polí	14
1.2 Lokalizace směru přicházejícího zvuku	15
1.2.1 Odhad úhlu přicházejícího zvuku pomocí 2 mikrofonů	15
1.2.2 Směr příchozího signálu	15
1.2.3 Časový rozdíl příchodu signálu	15
1.2.4 Frekvenční rozdílnost příchozího signálu	16
1.2.5 Metody založené na analýze energie signálu	17
1.3 Metody hledající směr příchozího zvuku	18
1.3.1 Metoda maximální entropie	18
1.3.2 Křížová korelace	19
1.3.3 Všeobecná křížová korelace	19
1.3.4 Transformace fáze	20
1.3.5 Pravděpodobnostní maximum	21
1.3.6 Metoda AAR	22
1.3.7 Nejbližší bod příchodu signálu	23
1.3.8 Metoda MUSIC	23
1.3.9 Určování polohy zvuku pomocí analýzy energického spektra	24
1.3.10 Energická metoda s STFT	26
1.3.11 Metoda Beamforming	28
1.3.12 Algoritmus ESPRIT	30
1.4 Matematické funkce	31
1.4.1 Diskrétní Fourierova transformace (DFT)	31
1.4.2 Krátkodobá Fourierova transformace (STFT)	31
1.4.3 Vzájemná korelace	33
<b>2 Vyhodnocení praktické části</b>	<b>34</b>
2.1 Popis měření a zpracování výsledků	34
2.1.1 Parametry měření s vyrobeným mikrofonním polem	35
2.1.2 Zpracování v Matlabu	38

2.2	Výsledky . . . . .	46
2.2.1	Výsledky zkušebního signálu . . . . .	46
2.2.2	Výsledky z měření v laboratoři s reálným signálem . . . . .	49
2.2.3	Výsledky z měření v reálném prostředí . . . . .	62
<b>3</b>	<b>Závěr</b>	<b>65</b>
	<b>Literatura</b>	<b>67</b>
	<b>Seznam příloh</b>	<b>71</b>
<b>A</b>	<b>Zdrojové výpisy z Matlabu</b>	<b>72</b>
<b>B</b>	<b>Obsah přiloženého CD</b>	<b>77</b>
	<b>Seznam symbolů, veličin a zkratk</b>	<b>78</b>

# SEZNAM OBRÁZKŮ

1.1	Ukázka nastavení mikrofونů pro B-formát. . . . .	25
2.1	Původní navržený držák mikrofonního pole [1] . . . . .	35
2.2	Vylepšené navržený držák mikrofonního pole . . . . .	36
2.3	Schéma rozestavení mikrofونů v křížovém poli. [1] . . . . .	37
2.4	Schéma rozestavení mikrofونů v kruhovém poli. [1] . . . . .	38
2.5	Ukázka nastavení mikrofونů pro B-formát, zkušební verze. . . . .	42
2.6	Graf míry korelace zkušebního signálu (ukázka)[1] . . . . .	48
2.7	Tabulky výsledků měření v laboratoři, přehled jednotlivých metod [1] . . . . .	51
2.8	Výsledky naměřeného úhlu metodami GCC, PHAT a ML v kruhovém a křížovém poli. Typ signálu: zabouchnutí dveří. . . . .	52
2.9	Výsledky naměřeného úhlu metodami GCC, PHAT a ML v kruhovém a křížovém poli. Typ signálu: rozbití skla. . . . .	53
2.10	Výsledky naměřeného úhlu metodami GCC, PHAT a ML v kruhovém a křížovém poli. Typ signálu: výstřel. . . . .	54
2.11	Výsledky průměrů naměřeného úhlu v křížovém poli. Typ signálu: dveře. Velikost signálu 1024 vzorků, úhel počítán na 4 vzorkách . . . . .	55
2.12	Výsledky průměrů naměřeného úhlu v křížovém poli. Typ signálu: sklo. Velikost signálu 1024 vzorků, úhel počítán na 4 vzorkách. . . . .	56
2.13	Výsledky průměrů naměřeného úhlu v křížovém poli. Typ signálu: výstřel. Velikost signálu 1024 vzorků, úhel počítán na 4 vzorkách. . . . .	57
2.14	Výsledky průměrů naměřeného úhlu v kruhovém poli. Typ signálu: dveře. Velikost signálu 1024 vzorků, úhel počítán na 4 vzorkách . . . . .	58
2.15	Výsledky průměrů naměřeného úhlu v kruhovém poli. Typ signálu: rozbití skla. Velikost signálu 1024 vzorků, úhel počítán na 4 vzorkách . . . . .	59
2.16	Výsledky průměrů naměřeného úhlu v kruhovém poli. Typ signálu: výstřel. Velikost signálu 1024 vzorků, úhel počítán na 4 vzorkách . . . . .	60
2.17	Četnosti nalezených úhlů výsledné matice energetické metody, typ signálu: růžový šum, A-formát signálu. Shora: pohyb o 10°, 20°, 30°, 40°, 50°, 60°, 70° . . . . .	61
2.18	Výsledky průměrů naměřeného úhlu v křížovém poli a reálném pro- středí. Typ signálu: siréna. Velikost signálu 1024 vzorků, úhel počítán na 4 vzorkách. . . . .	63
2.19	Výsledky průměrů naměřeného úhlu v křížovém poli a reálném pro- středí. Typ signálu: výstřel. Velikost signálu 1024 vzorků, úhel počítán na 4 vzorkách. . . . .	63

2.20	Výsledky průměrů naměřeného úhlu v křížovém poli a reálném prostředí. Typ signálu: rozbití skla. Velikost signálu 1024 vzorků, úhel počítán na 4 vzorkách. . . . .	64
------	--	----

# SEZNAM TABULEK

2.1	Posun úhlu v závislosti na zpoždění při vzorkovacím kmitočtu 48 kHz.	
	[1] . . . . .	46
2.2	Posun úhlu v závislosti na zpoždění při vzorkovacím kmitočtu 96 kHz.	
	[1] . . . . .	47
2.3	Přehled jednotlivých mikrofonů v poli. . . . .	49
2.4	Výsledky měření v reálném prostředí, přehled jednotlivých metod. . .	62

## SEZNAM VÝPISŮ

A.1	GCC metoda naprogramovaná v Matlabu. . . . .	72
A.2	Část kódu naprogramované metody PHAT v Matlabu. . . . .	73
A.3	Část kódu naprogramované metody ML v Matlabu. . . . .	73
A.4	Jednoduché naprogramování umělého zvuku v Matlabu. . . . .	73
A.5	Energetická metoda naprogramovaná v Matlabu . . . . .	74
A.6	Energetická metoda naprogramovaná v Matlabu . . . . .	75
A.7	Energetická metoda naprogramovaná v Matlabu . . . . .	76



# ÚVOD

Odhad směru příchodu zvuku byl a je velkou výzvou pro digitální zpracování signálu. Jednou z možností, jak najít správný směr, je využít mikrofonní pole. Na toto téma je napsáno mnoho prací a bylo vynalezeno mnoho metod, které se vzájemně mohou doplňovat a překrývat. Tyto metody se dají aplikovat nejen na zvukové vlnění, ale samozřejmě i na elektromagnetické.

Tato diplomová práce navazuje na bakalářskou práci stejného názvu, rozšiřuje ji o nové další metody. Je zde přehled metod, které využívají časovou, frekvenční i energetickou analýzu jako například CPA, GCC, GCC-PHAT, MUSIC, beamforming a další. Také je zde popsána potřebná matematická i fyzikální teorie k pochopení jednotlivých metod. V praktické části jsou okomentované naprogramované funkce v programu Matlab a určena jejich přesnost a spolehlivost v závislosti na frekvenčním rozložení signálu a typu pole. Zejména jsem se zaměřil na hledání příchodu směru zvuku pomocí malých vzorků na malé délce signálu. Je zde také realizováno potlačení ruchů.

Bohužel možnost vyzkoušet nově navržené pole byla narušena pandemickou krizí, a tím tak možnost získání nových nahrávek, jak v prostředí laboratoře, tak v reálném prostředí. Vycházím tedy z dat naměřených při bakalářské práci.

# 1 TEORETICKÁ ČÁST PRÁCE

## 1.1 Zvuk

### 1.1.1 Šíření zvuku prostorem

Zvuk je v podstatě kmitání molekul v prostoru, nebo-li kmitání hmotného prostředí. Vlastnosti jeho šíření závisí na jednotlivých vazbách molekul v prostředí (ve vzduchu se zvuk šíří pomaleji, protože jsou molekuly vzduchu dál od sebe než například ve vodě), tudíž třeba ve vakuu nemůže zvuk ani existovat. Když je prostor neomezený a bezodrazový, označujeme jej za volný. [1]

#### Rychlost šíření zvuku ve vzduchu

Za předpokladu, že suchý vzduch má barometrický tlak, tj.  $p_{00} = 1013 \text{ kPa}$ , potom rychlost zvuku ve vzduchu o teplotě  $t$  (ve  $^{\circ}\text{C}$ ) se dá jednoduše spočítat jako [1] [2]

$$v = (331,6 + 0,61t) \text{ m} \cdot \text{s}^{-1}. \quad (1.1)$$

V mé práci budu počítat s hodnotou  $344 \text{ m} \cdot \text{s}^{-1}$ .

### 1.1.2 Akustický tlak

Je brán jako rozdíl mezi okamžitou výchylkou celkového tlaku a vztažnou referenční hodnotou  $p_0$ , což je barometrický (atmosferický) tlak. Kvůli kmitání vlny (pro zjednodušení můžeme vzít funkci sinus) se tedy tlak mění v čase za přítomnosti zvuku. Tyto výchylky poté můžeme zachytit na mikrofon, což je ve své podstatě měnič výchylky akustického tlaku na elektrickou vlnu. [1]

### 1.1.3 Akustická impedance

Šíření tlaku (zvuku) má za následek rozkmitání prostředí. Kinetická energie částic a potenciální energie prostředí se přelévá, což se dá analogicky popsat jako pohyb nábojů v elektrickém poli. Pro definování speciální akustické impedance je odvozen vzorec:[2]

$$z_a = \frac{p}{v} = \rho \cdot c_0, \quad (1.2)$$

kde  $p$  je akustický tlak v místě a  $v$  je akustická rychlost. Díky akustické impedanci se na rozhraní dvou prostředí může zvuk lámat nebo odrážet. Jeho velikost závisí na mnoha veličinách (atmosferický tlak, teplota vzduchu, hustota prostředí). V mé práci budu brát velikost akustické impedance vzduchu  $z_A = 413 \text{ kg} \cdot \text{m}^2 \cdot \text{s}^{-1}$ . [2]

### 1.1.4 Typy zvukových polí

Jak již bylo naznačeno, existuje tzv. volné pole, kde neexistují odrazy. Pokud je ve směru zvuku nějaká překážka, musí zákonitě dojít k odrazu a pohlcení energie, případně k ohybu zvukové vlny. V uzavřeném prostoru poté dochází k interferenci odražených (sekundárních) a přímých (primárních) vln a pokud se tento děj opakuje, tak akustické pole odražených vln nazýváme difúzní pole. V difúzním poli nedokážeme určit směr šíření zvukové energie, jelikož je náhodný.

Za předpokladu, že sekundární vlny v difúzním poli mají malou energii, můžeme toto pole zanedbat a považovat prostor za volné pole. Tento předpoklad jsem využil právě při pokusech v bezodrazové komoře. Dalším rozřazením je vzdálené pole. Ve vzdáleném poli vzdálenost zdroje zvuku od mikrofonu či pole mikrofونů je srovnatelná s velikostí tohoto zdroje. Potom se vlna považuje za rovinnou a nemusí se brát v potaz zakřivení vlnoplochy. [1] [2]

### 1.1.5 Typy mikrofonních polí

Mikrofonní pole malých rozměrů má řadu omezení při jeho návrhu. Zejména je potřeba vytvořit pole uniformní (má stejné vlastnosti ve všech směrech), aby se dalo využít na více než jenom jednu metodu zpracování signálu. Každá metoda má svá vlastní omezení, které jsou hlavně ve složitosti výpočtu. Jsou zde i omezení fyzikální, a to nejvyšší možná frekvence snímaná bez interference a zkreslení, ale zároveň i nižší. Některé algoritmy nedokáží zpracovat vlnové délky větší než je velikost pole. Jsou to hlavně takové, které určují směr příchozího zvuku pomocí rozdílu fáze. Je nutné, aby fyzické rozměry pole (v mém případě plastová destička), měly co nejmenší dopad na přijímaný zvuk (opět možnost odrazové plochy pro zvuk či vytvoření ruchu). Je možné také pole přizpůsobit akustickému prostředí, nicméně mnoho algoritmů má již implementovanou filtraci rušivých zvuků nebo dokonce dokáže rozeznat více zdrojů. Geometrie pole je nedílnou součástí řešení problému. Mohou být pole pravidelná a nepravidelná, lineární, rovinná, prostorová. Liší se především tím, že lineární se lépe matematicky popisují, ale nepravidelná mají výhodu možnosti více kombinací vzdáleností. V praxi se poté používají například lineární, křížová, kruhová, logaritmicky zakřivená pole. [1] [15]

## 1.2 Lokalizace směru přicházejícího zvuku

Odhad úhlu přicházejícího signálu (AOA–Angle of Arrival) je proces, který vymezuje směr příchozího signálu skrze zpracování signálu na snímač (mikrofon, senzor, anténu). Využívá se v mnoha oborech jako vojenství, ale i v civilních odvětvích.

### 1.2.1 Odhad úhlu přicházejícího zvuku pomocí 2 mikrofonů

Pomocí dvou mikrofonů (ležící v prostoru na osách  $x, y$ ) dokážeme lehce určit úhel přicházejícího zvuku, avšak již nedokážeme určit, jestli zvukový zdroj leží mimo plochu  $x, y$  (jestli jde zvuk zdola či shora). Abychom dokázali určit posun na ose  $z$ , tak mikrofony musí ležet na osách  $x, y, z$ . Z toho plyne, že k detekci úhlu v prostoru je nutné použít minimálně tři mikrofony.

Úhel přicházejícího signálu vůči poli jednoduše podle můžeme spočítat jako [1]

$$\theta = \arccos \frac{\Delta t \cdot c}{d}. \quad (1.3)$$

Proměnná  $d$  je vzdálenost mezi dvěma měřícími mikrofony,  $\Delta t$  je časové zpoždění zvuku mezi mikrofony a  $c$  je rychlost zvuku při dané teplotě vzduchu (vlhkost a tlak je zanedbán).

### 1.2.2 Směr příchozího signálu

Směr příchozího signálu (DOA–Direction of Arrival) je termín, který se používá v signálovém zpracování a značí směr vln přicházejícího zvuku k bodu, kde je umístěn (většinou referenční) mikrofon či snímač. K určení směru příchodu používáme mikrofonní pole (pole senzorů). [17]

### 1.2.3 Časový rozdíl příchodu signálu

Časový rozdíl příchodu signálu (TDOA–Time Difference of Arrival) je běžnou pasivní metodou pro zjišťování směru příchodu zvuku. Signál snímáný senzorem můžeme definovat pro TDOA jako [16]:

$$x_1(t) = s(t) + \xi_1(t), \quad (1.4)$$

$$x_i(t) = \frac{1}{g_{i1}} s(t - \tau_{i1}') + \xi_i(t), \quad (1.5)$$

kde  $i = 2, 3 \dots M$ . Příchozí signál je  $s(t)$ ,  $x_i(t)$  je šum přijímaný na senzoru  $i$ ,  $x_i(t)$  je výsledný signál.  $\tau_{i1}'$  je definován jako časový rozdíl mezi mikrofony o počtu  $M$ . Jednotlivá zpoždění jsou potom  $\tau' = [\tau_{21}', \tau_{31}', \dots, \tau_{M1}']^T$ . Výsledkem je, že rozdíl vzdálenosti příchozího signálu (DDOA–Distance Difference of Arrival) na  $i$  senzoru

a rozdíl mezi jedním snímačem je nutné vynásobit rychlostí signálu  $c$ . Vzdálenost mezi  $i$  snímačem a cílem může být zapsáno jako: [16]:

$$r_i' = \|u' - s_i'\|, \quad (1.6)$$

přičemž  $\|*\|$  znamená euklidovskou vzdálenost,  $i = 2, 3 \dots M$ . Vzdálenost mezi  $i$  senzorem a prvním je potom:

$$r_{i1}' = r_i' - r_1' = c\tau_{i1}' = r_{i1} + \Delta r_{i1}. \quad (1.7)$$

Proměnná  $\tau_{i1}'$  je časový rozdíl mezi prvním a  $i$  snímačem,  $r_{i1}$  je dopočítaná vzdálenost a  $\Delta r_{i1}$  je chyba měření vzdálenosti.

Vyjádřeno vektory,  $r' = [r_{21}', r_{31}', \dots, r_{M1}']^T$ ,  $r = [r_{21}, r_{31}, \dots, r_{M1}]^T$ ,  $\Delta r = [\Delta r_{21}, \Delta r_{31}, \dots, \Delta r_{M1}]^T$ . Potom rovnice vzdáleností je: [16]:

$$r = r' + \Delta r \quad (1.8)$$

Předpokládaná chyba využívá Gaussovo rozložení a kovarianční matice je  $E[\Delta r \Delta r^T] = Q_r$ .

### 1.2.4 Frekvenční rozdílnost příchozího signálu

Frekvenční rozdílnost příchozího signálu (FDOA–Frequency of Arrival/Differential Doppler) využívá posunutí frekvencí signálu, které vzniká Dopplerovým efektem (objekt nebo pole se musí hýbat). Měří, jak moc se liší frekvence v signálu na jednotlivých mikrofonech. Problémem hledání směru pomocí FDOA je takový, že výsledné funkce jsou nelineární, a tím pádem jsou lehce ovlivnitelné šumem.

Za předpokladu, že zdroj hluku je neměnný, posunutí signálu způsobené Dopplerovým efektem na mikrofonu  $i$  je: [18]

$$d_i = \frac{f_0}{c} \left( v_i \cdot \frac{x - x_i}{\|x - x_i\|} \right). \quad (1.9)$$

Přičemž

$$v_i = (u_i, v_i, w_i)^T \quad (1.10)$$

obsahuje jednotkový vektor ukazující na směr zdroje zvuku. Tento výraz je brán ke střední frekvenci signálu ( $f_0$ ), dělená rychlostí zvuku  $c$ . V praxi často  $f_0$  je neznámá, proto se ponechává jako proměnná. S výjimkou  $\frac{f_0}{c}$  je vzdálenost  $d_i$  úměrná k frekvenčnímu posunu signálu na mikrofonu  $i$ . Potom vztah FDOA může být vyjádřen takto:[18]

$$f_{1,i} = d_i - d_1 = v_i \cdot \frac{x - x_i}{\|x - x_i\|} - v_1 \cdot \frac{x_1 - x}{\|x_1 - x\|}. \quad (1.11)$$

Aby bylo zajištěno neovlivnění šumem, musí být splněna následující podmínka

$$f_{i,j} + f_{j,k} - f_{i,k} = 0. \quad (1.12)$$

Nelinearita FDOA je problematická i v tom, že může mít více řešení. Proto se FDOA řeší jako systém polynomičských rovnic. Pro zjištění oblasti polohy zdroje zvuku je nutné znát rychlost a polohu přijímačů. Následné vyřešení souboru polynomičských rovnic zjistí správnou polohu zdroje. FDOA se hodí zvláště pro určování směru zvuku, který má úzké spektrum a dlouhé trvání.

### 1.2.5 Metody založené na analýze energie signálu

#### Metoda GROA

Algoritmus rozsahu vstupních úrovní přicházejícího signálu (GROA–Gain Ratios of Arrival) byl poprvé navržen v studiích sepsaných autory K. C. Ho a S. Ming.[16] Nejprve byl spřažen, a to v metodě váhování nejmenšími čtverci (TSWLS–Weighted Least Squares) k měření TDOA, následně vylepšován. V rovnici (1.5) je  $g_n$  úroveň vstupního signálu na senzoru  $i$  vztažená k prvnímu senzoru. Metoda GROA je postavena na faktu, že energie zvuku šířící se ve volném prostoru klesá se čtvercem vzdálenosti. Úroveň vstupního signálu jednotlivých mikrofónů  $M$  je: [16]

$$g' = [g_{21}', g_{31}', \dots, g_{M1}']^T. \quad (1.13)$$

Potom GROA mezi  $i$  senzorem a prvním může být vyjádřeno vztahem

$$g_{i1}' = \frac{r_i'}{r_1'}, \quad (1.14)$$

kde  $r_i'$  je vzdálenost mezi cílem a  $i$  senzorem. Parametry se dají zapsat jako vektor  $g = [g_{21}, g_{31}, \dots, g_{M1}]^T$ ,  $\Delta g = [\Delta g_{21}, \Delta g_{31}, \dots, \Delta g_{M1}]^T$ . Dále můžeme napsat rovnici GROA jako: [16]

$$g = g_{i1}' + \Delta g_{i1}, \quad (1.15)$$

přičemž  $g_{i1}$  je reálná hodnota GROA,  $\Delta g_{i1}$  je chyba měření GROA.

#### Modelování poklesu energie

Metoda modelování poklesu energie opět využívá toho (podobně jako GROA), že pokud se zvuk šíří vzduchem, tak akustická energie vyzářená všemi směry od zdroje se zmenšuje se čtvercem vzdálenosti. Energie signálu  $y_i(t)$  naměřená na mikrofónu  $i$  v časovém intervalu  $t$  můžeme vyjádřit jako: [19]

$$y_i(t) = g_i \cdot \frac{s(t - t_i)}{|r(t - t_i) - r_i|^\alpha} + \epsilon_i(t), \quad (1.16)$$

kde  $t_i$  je časové zpoždění signálu od zdroje k mikrofonu  $i$ ,  $s(t)$  je skalární veličina, která vyjadřuje energii signálu v časovém intervalu  $t$ ,  $r(t)$  je vektor  $d \times 1$  obsahující souřadnice pohybu zdroje zvuku v čase  $t$ ,  $r_i$  je vektor  $d \times 1$  značící kartézské souřadnice polohy mikrofonu  $i$  a  $\alpha (\approx 2)$  je faktor poklesu energie, potom  $\epsilon_i(t)$  je efekt modelované chyby, který se může nahradit kladnou střední hodnotou.

Vzhledem k maximální velikosti pole (20 cm) by bylo nejvíce vhodné kruhové pole mikrofonů. Nicméně i tak je pokles energie na takovou vzdálenost velmi malý, proto tato metoda není vhodná pro malá pole.

## Metoda analýzy energie

Tato metoda je postavena na principu toho, že zdroj zvuku leží na opačné straně vektoru intenzity zvuku. Energie zvuku je rozložena v čase a frekvencích, která lze ukázat pomocí spektrogramu. Při předpokladu, že je více zdrojů zvuku, je možné, že pokud budou mít podobné energeticko-frekvenční složení, nebude možné najít směr příchozu zvuku.

## 1.3 Metody hledající směr příchozího zvuku

### 1.3.1 Metoda maximální entropie

Metodu maximální entropie (MEM–Maximum entropy method) navrhl John Burg v roce 1967 jako alternativu k metodám využívající spektrum signálu, které zanedbávaly zpoždění korelace. Ačkoli to nebylo na první pohled jasné, Burg obhajoval, že výsledky na základě spektrální analýzy mohou být odvozeny bez autoregresivních funkcí v časové oblasti. MEM je odvozeno z lineární predikce chyby filtru. První koeficient takového filtru je jednotný, zbylé koeficienty jsou voleny tak, aby minimalizovaly svůj účinek na výstup s ohledem na predikci chyby. Návrh jednotlivých koeficientů filtru by vypadal následovně: [20]

$$Rw = [e, 0 \dots 0]^T \quad (1.17)$$

pro předpokládanou chybu  $e$  a neznámé prvky matice

$$w = [1, ? \dots ?]^T, \quad (1.18)$$

přičemž  $R$  je teoretická kovariační matice prvků. Převedením do spektrální oblasti dostaneme rovnici, která může být napsána jako: [20]

$$\hat{S} = \frac{1}{|v^H(\alpha)R^{-1}u_1|^2}, \quad (1.19)$$

kde  $u_1$  je první sloupec jednotkové matice. Tyto výrazy mohou být takto interpretovány jako maximální spektrální hustota entropie pouze v ideálním případě, že vstupní (chtěný) signál nekoreluje s jiným ruchem. Další podmínkou je jednotné lineární pole (stejné rozestupy mezi mikrofony, které jsou 0. řádu). Následné vyhlazování filtru a nalézání koeficientů může být prováděno odhadem, tak, aby vyšla Toeplitzova struktura (například Levinsonovou rekurzí či Yule-Walkerovou metodou). [20] MEM implementace je výhodná a účinná zejména kvůli minimální změně fáze filtru.

### 1.3.2 Křížová korelace

Je hned několik metod, jak zjistit časový rozdíl mezi dvěma signály, avšak nejjednodušší metodou je křížová korelace.[3] Jestliže známe dva průběhy signálů, z nichž jeden z nich je zpožděný, potom nalezením maxima jejich křížové korelace získáme časový posun signálu mezi mikrofony. Zobecněním rovnic (1.4) pro více signálů vyplývá tvar [3]

$$x_i = a_i s_i(t + \tau_i) + n_i(t), \quad (1.20)$$

kde  $i$  v indexu značí, ze kterého mikrofonu je brán signál  $s(t)$ ,  $\tau_i$  je hledaný časový rozdíl mezi zvolenými mikrofony,  $a_i$  označuje útlum signálu. V rovnici je přidán šum  $n(t)$ , který musí být na signálu nezávislý, jinak by mohl být považován za signál, což by mohlo vést ke znevážení výsledků. Podrobněji se touto závislostí zabývá G. C. Carter [4]. A právě kvůli šumu je v mnoha metodách hledání  $\tau_i$  (transformace fáze, pravděpodobnostní maximum) signál filtrován váhovací funkcí. Pro snadnější nalezení  $\tau_i$ , je vhodné provést Fourierovu transformaci signálu, která je více popsána v sekci matematických funkcí (1.4). Vyplyne z tohoto kroku rovnice ve tvaru [3]

$$X_i(f) = \alpha_i S_i(f) \cdot e^{-i2\pi f \tau_i} + N_i(f), \quad (1.21)$$

kde  $f$  je frekvence. Potom výsledná obecná korelační funkce vyjde [3], [5]

$$R_{x_1 x_2} = E[x_1(t) \cdot x_2(t - \tau)]. \quad (1.22)$$

$E[.]$  je předpokládaná funkce, která se liší podle toho, jak chceme signál filtrovat. Pokud za  $E = 1$ , potom dostaneme křížovou korelaci v časové doméně. Příkladem jiné funkce může být třeba tato rovnice [5]

$$R_{x_1 x_2}(t) = \frac{1}{\tau - T} \int_{\tau}^T x_1(t) \cdot x_2(t - \tau) dt. \quad (1.23)$$

### 1.3.3 Všeobecná křížová korelace

Všeobecná křížová korelace (GCC–Generalized Cross–Correlation) je vylepšená verze křížové korelace. Největší výhodou tohoto algoritmu je relativní přesnost a malé nároky na výpočetní funkci. Jak už bylo zmíněno, nejčastěji se používá signál, který již



prošel přes číslicový filtr, a to kvůli tomu, aby výsledná korelace měla jednoznačné maximum a byla zajištěna vyšší přesnost výsledku. Kromě číslicových filtrů tuto skutečnost zajišťuje například lepší odstup signálu od šumu (SNR), správně zvolený vzorkovací kmitočet, analogové filtry vřazené do signálové cesty mezi mikrofon a samotný proces nahrávání. [3], [5], [4]

$$S_{x_1x_2} = E[X_1(f) \cdot X_2^*(f)] \quad (1.24)$$

Tato rovnice je výpočet křížového spektra mezi signály  $X_1(f)$ ,  $X_2(f)$ . Předpokládáná funkce  $E[.]$  má stejný význam jako v (1.22) a  $(.)^*$  značí komplexně sdružené číslo. Rovnice všeobecné křížové korelace v tomto případě potom bude [3]

$$\Psi_{x_1x_2}(f) = \int_{-\infty}^{\infty} \phi(f) S_{x_1x_2}(f) e^{i2\pi f\tau} dt. \quad (1.25)$$

V této rovnici  $\phi(f)$  značí kmitočtovou váhovací funkci, což může být například impulzní odezva prostoru. Integrál v praxi nahrazujeme zpětnou Fourierovou transformací. Časové zpoždění  $\tau_\phi$  získáme tak, že najdeme maximum této rovnice, matematicky zapsáno [3], [5], [4]

$$\tau_\phi = \arg \max \Psi_{x_1x_2}(f). \quad (1.26)$$

Výběrem váhovacího filtru v rovnici (1.25) značně ovlivníme výsledek a je jasné, že čím složitější funkci zvolíme, tím větší nároky klademe na výpočetní techniku. Pro všeobecnou křížovou korelaci bychom zvolili  $\phi(f) = 1$ . Další metody určování časového rozdílu, například transformace fáze nebo pravděpodobnostní maximum, jsou odvozeny právě z křížové korelace a liší se pouze ve váhovacím filtru. Pouhá křížová korelace je velice účinná v prostředí, kde není velký dozvuk (menší než 0,15 s) a signál není rušen šumem či hlukem. [3]

### 1.3.4 Transformace fáze

Základem této metody je, že vezme křížové spektrum signálů  $x_1$ ,  $x_2$  a převede jej na konstantu. Váhovací funkci transformace fáze můžeme vyjádřit jako [3]

$$\psi_p(f) = \frac{1}{|G_{x_1x_2}(f)|}, \quad (1.27)$$

dosazením do (1.25) potom získáme [3]

$$R_p(\tau) = \int_{-\infty}^{\infty} \frac{1}{|G_{x_1x_2}(f)|} G_{x_1x_2}(f) e^{j\omega\tau} dt. \quad (1.28)$$

Výsledné časové zpoždění vyjádříme ve tvaru [3]

$$\hat{\tau}_p = \arg \max_{\tau \in D} R_p(\tau). \quad (1.29)$$

Tato funkce zachovává fázi signálů, proto se dá použít i na kontrolování polarity. Hlavní výhodou je, že vylučuje z křížového spektra energii signálu, která v ostatních filtrech může ovlivnit výsledek.

### 1.3.5 Pravděpodobnostní maximum

Pravděpodobnostní maximum (Maximum Likelihood–ML) je metodou, která využívá akustickou energii signálu na jednotlivých senzorech zvlášť, díky tomu nevyžaduje geometricky symetrické pole. Může být využívána pro lokalizaci více zdrojů zvuku. [21] Základní myšlenka je jednoduchá – do každého mikrofону dojde signál v jiném čase. Potom váhování každého signálu s jistým zesílením a s ohledem na daný úhel dopadajícího signálu. Musí se také minimalizovat zesílení signálu polem. Proto se výstupní signál z takového pole neovlivňuje. Pokud je ruch Gausovského typu, potom výstupní energie je nazývána jako odhad maximum likelihood ze směru  $\alpha$ .

Matematicky se následně dá tento vztah zapsat následovně:[20]

$$P = w^H \cdot \hat{R} \cdot w. \quad (1.30)$$

Proměnná  $w$  zde je váhování. Následným krokem je omezení zesílení pole, tak aby bylo jednotné ve směru  $\alpha$ .

$$w^H v(\alpha) = 1 \quad (1.31)$$

Minimalizování může být provedeno pomocí metody Lagrangových multiplikátorů. Optimální váhování (maximizace výstupu SIR–Signal-to-Interference Ratio ) lze najít poté vyřešením následující rovnice: [20]

$$\hat{R} \cdot w(\alpha) = \lambda \cdot v(\alpha), \quad (1.32)$$

kde Lagrangianova lambda je volena na základě (1.31). Pronásobením této rovnice konjugovanou maticí  $w(\alpha)$ , je možné zjistit, že MLM a Lagrangovy multiplikátory vychází číselně stejně. Nahrazením  $\lambda$  proměnnou  $P$ , řešením  $w(\alpha)$ , a dosazením výsledku do (1.30) vede k MLM odhadu [20]

$$\hat{P}_{\text{MLM}} = \frac{1}{v^H(\alpha) \hat{R}^{-1} v(\alpha)}. \quad (1.33)$$

Pravděpodobnostní maximum lze vyjádřit i pomocí křížové korelace. Metodu je potom možné odvodit z rovnice (1.25) dosazením tohoto filtru [5]

$$\psi_{\text{ML}}(f) = \frac{1}{|G_{x_1 x_2}(f)|} \cdot \frac{|\gamma_{x_1 x_2}(f)|^2}{1 - |\gamma_{x_1 x_2}(f)|^2}, \quad (1.34)$$

kde

$$R_{\text{ML}}(\tau) = \int_{-\infty}^{\infty} \psi_{\text{ML}}(f) \cdot X_1(f) \cdot X_2^*(f) \cdot e^{i2\pi f\tau} df = \int_{-\infty}^{\infty} \psi_{\text{ML}}(f) \cdot G_{x_1x_2}(f) \cdot e^{i2\pi f\tau} df, \quad (1.35)$$

$\psi_{\text{ML}}$  je váhovací funkcí,  $X_i(f)$  značí Fourierovu transformaci  $x_i(f)$ ,  $X_1^*(f)$  je komplexně sdružené číslo k  $X_1(f)$ ,  $G_{x_1x_2}(f)$  představuje křížové spektrum přijímaného signálu. Abychom mohli spočítat pravděpodobnostní maximum (jeho váhovací funkci (1.34)), zjišťujeme míru koherence vstupních dat, což znamená, do jaké míry může vycházet signál  $x_1$  ze signálu  $x_2$ . [5], [3]

$$|\gamma_{x_1x_2}(f)|^2 = \frac{|G_{x_1x_2}(f)|^2}{G_{x_1x_1}(f)G_{x_2x_2}(f)}. \quad (1.36)$$

Rovnice využívá metodu nejmenších čtverců, což oproti PHAT (1.3.4) by mělo být přesnější, protože křivka prokládající spektrum není pouhou konstantou, ale snaží se jej dle nějaké křivky kopírovat. Časové zpoždění můžeme zjistit pomocí maxima [3]

$$\hat{\tau}_{\text{ML}} = \arg \max_{\tau \in D} R_{\text{ML}}(\tau). \quad (1.37)$$

Tato metoda je díky tomu, že korelaci provádí ve spektrálních složkách (kde je lepší odstup signálu od šumu), velmi přesná. Její nedostatek je v tom, že na určitých frekvencích interferuje a může tak tvořit lokální maxima. [9]

### 1.3.6 Metoda AAR

Metoda odezvy úhlových kmitočtů (AAR–Adapted Angular Response) je algoritmus, který navrhli G. V. Borgiotti a L. J. Kaplan a jde vlastně o variaci na MLM. Má stejnou matematickou strukturu jako ARMA (Autoregressive Moving Average Model) pro odhad výkonové spektrální hustoty. Upravuje amplitudu vstupního signálu tak, že součet druhých mocnin amplitudy je konstantou např. [20]

$$|w(\alpha)|^2 = 1 \quad (1.38)$$

Tato modifikace vede k metodě hledání směru jejímž výsledkem je, že efekt bílého šumu je v průměru rozložen do všech směrů stejně. Vybráním Lagrangianové lambdy, tak, aby platilo (1.38), vede k AAR odhadu energie rovnicí [20]

$$\hat{P}_{\text{AAR}} = \frac{v^H(\alpha) \hat{R}^{-1} v(\alpha)}{v^H(\alpha) \hat{R}^{-2} v(\alpha)}. \quad (1.39)$$

### 1.3.7 Nejbližší bod příchodu signálu

Název této metody (CPA–Closest Point of Aproach), je orientační a charakterizuje nejbližší bod dvou objektů pohybujících se po nekřížících se přímkách. Metoda, která dokáže nalézt takový bod, je taková, která najde mikrofon s největší naměřenou energií:[21]

$$i* = \arg \max y_i. \quad (1.40)$$

Tím pádem je možné stanovit místo, odkud zdroj zvuku přichází jako místo senzoru:

$$\rho = r_{i*} \quad (1.41)$$

Takto se metoda používá pouze pro jeden zdroj zvuku  $r_{i*}$ . Pokud je zdrojů více, algoritmus musí najít všechna lokální maxima všech mikrofonů  $r_i$ . Zjednodušeně vyjádřeno: [21]

$$\rho_k = \frac{\left(\sum_{i=1}^N \alpha_{ik} r_i\right)}{\left(\sum_{i=1}^N \alpha_{ik}\right)} \quad (1.42)$$

V této rovnici je  $\rho_k$  zdroj zvuku v pořadí  $k$ ,  $\alpha_{ik}$  je směr příchozího zvuku, který se dá vyjádřit rozvinutím takto: [21]

$$\alpha_{ik} = \frac{g_i}{\sigma_i} \left( \frac{1}{d_{ik}^4} \right) \left( z_i - \frac{g_i}{\sigma_i} \sum_{m=1}^K \frac{S_m}{d_{im}^2} \right) \quad (1.43)$$

Tato rovnice popisuje ML, pro získání matematického popisu CPA je třeba ji zjednodušit. Pokud  $K = 1$  a  $d_i = d_{ik}$ , potom úhel směru získaný CPA je [21]

$$\alpha_i = \frac{g_i}{\sigma_i} \left( \frac{\xi_i}{d_i^4} \right) \quad (1.44)$$

za předpokladu, že  $g_i, \sigma_i$ , a  $\xi_i$  mají stejný řád pro všechny senzory.  $\xi_i$  je možné vyjádřit jako náhodná čísla s gausovským rozložením  $\xi_i = (\epsilon_i - \mu_i)/\sigma_i \sim N(0, 1)$ . Více o CPA je zde [22].

### 1.3.8 Metoda MUSIC

Algoritmus MUSIC (The Multiple Signal Classification) je postaven na metodě objevování harmonických složek signálu, kterou prvně popsal Pisarenko. [20] Základní myšlenkou algoritmu je, že počet přijímaných zdrojů zvuku je menší než počet přijímačů (mikrofonů). Potom kovarianční matice přijímaného signálu [20]

$$S = E \{ s s^H \} \quad (1.45)$$

je singulární. S přihlédnutím k (1.17) můžeme chápat tuto rovnici jako nástroj pro částečnou extrapolaci („predikci“) přijímaného signálu.

Prostor zahrnut ve sloupcích matice  $S$  je vztažen k prostoru signálů (soubor všech rovnic popisující příchozí signály). S výjimkou výjimečných situací, kdy jsou vstupní signály podobné (korelované), leží vektory směru příchozího zvuku v prostoru signálů. Pokud tomu tak není, rozhodující vektor směru je částečně obsažen v (doplňkovém) nulovém prostoru matice  $S$ .

Pokud je nulový prostor v orthonormálních souřadnicích, euklidovská vzdálenost od vektoru  $x$  k prostoru signálů se dá dopočítat projekcí  $x$  do nulového prostoru. Tato projekce se dá zapsat jako [20]

$$x_N = E_N E_N^H x, \quad (1.46)$$

kde sloupce matice  $E_N$  je orthonormální vektorová báze pro nulový prostor. Potom je vzdálenost od  $x$  k prostoru signálů

$$|x_N| = \left( E_N E_N^H x \right)^{\frac{1}{2}} \quad (1.47)$$

a

$$E_N^H E_N = I \quad (1.48)$$

je jednotková matice.

Rovnice pseudospektra metody MUSIC je potom inverze čtverce vzdálenosti od vektoru směru zvuku k prostoru signálů [20]

$$P_{\text{MUSIC}} = \frac{1}{v^H(\alpha) E_N E_N^H v(\alpha)}. \quad (1.49)$$

Teoreticky vrcholy spektra  $P_{\text{MUSIC}}$  vyjdou nekonečné, pokud je směr správný, nicméně kvůli chybám v nulových vektorech směru je nepravděpodobné, že by se tak stalo. A právě tyto maxima funkce je směr přicházejícího zvuku. Kvůli možnosti více maxim je potřeba parametricky zadat, kolik je hledaných počtů zvuku  $D$ . [23]

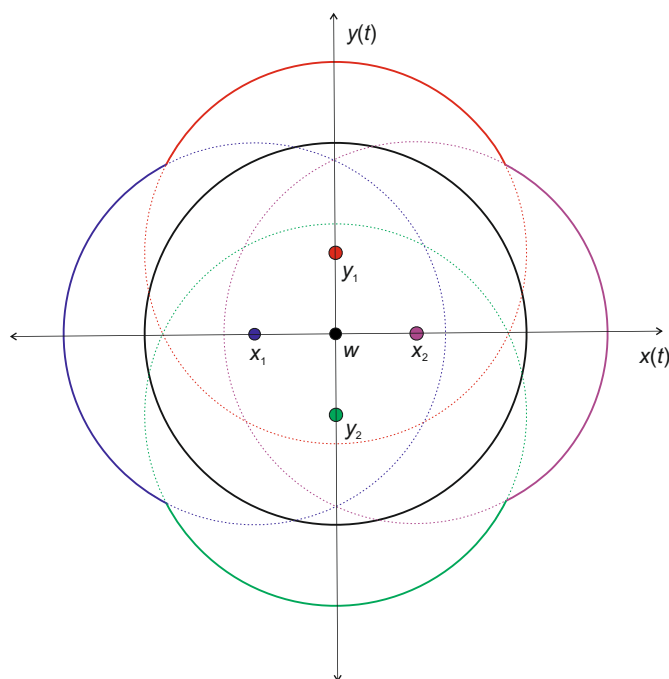
Pro lineární pole, jmenovatel pseudospektra je polynomická funkce v  $e^{ikd \cos \alpha}$  a může být také vyjádřen jako polynom v komplexní rovině. V takovém případě lze využít metod, které dokáží hledat kořeny,  $z_i$ . Ty nemusí ležet na jednotkové kružnici, avšak kořeny (o počtu  $D$ ) nejbližší kružnici jsou ty, které ukazují správný směr příchodu zvuku.

### 1.3.9 Určování polohy zvuku pomocí analýzy energického spektra

#### B-formát signálu

Tato metoda využívá B-formát signálu, který lze získat pomocí dvou mikrofónů 1. řádu (tzv. osmičkové) a jednoho mikrofónu 0. řádu (všesměrového). Dva mikrofóny

1. řádu lze teoreticky nahradit čtyřmi mikrofony 0. řádu ve čtvercové konfiguraci s co nejmenší délkou strany, potom signál je nutné ale odečíst  $p_{celkove} = p_{n_1} - p_{n_2}$ , kde  $p_1$  a  $p_2$  je tlak na jednotlivé membrány, tudíž jejich odečtení nám dá ve výsledku tlak  $\sum p$  na mikrofону 1. řádu. [2] V tomto formátu je možné určit, jak horizontální, tak i vertikální úhel. B-formát obsahuje čtyři signály:  $x(t)$ ,  $y(t)$ ,  $z(t)$ ,  $w(t)$ , ze kterých lze následně spočítat směr přicházejícího zvuku. Signály  $x(t)$  a  $y(t)$  dávají informace o horizontální rovině, jsou zaznamenávány právě na dva osmičkové mikrofony (čtyři všesměrové). [1]



Obr. 1.1: Ukázka nastavení mikrofونů pro B-formát. Jednotlivé dvojice leží na osách a pátý mikrofون ve středu souřadnic. Všechny mikrofony jsou kulového charakteru.[1]

$x(t)$  určuje předo-zadní rovinu a  $y(t)$  určuje levo-pravou rovinu. Signál  $z(t)$  reprezentuje vertikální osu a  $w(t)$  je referenční signál nahraný na všesměrový mikrofون, který však není nutný. Z goniometrie můžeme určit tento druh mikrofону ze čtyřech mikrofونů jako  $w = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$ , kde  $x_1$ ,  $x_2$  jsou dva mikrofony ležící na ose  $x$  a  $y_1$ ,  $y_2$  jsou dva mikrofony ležící na ose  $y$ . Při vzdálenosti mikrofونů od sebe zhruba do 7 mm se bude tento součet mikrofونů chovat přibližně jako jeden kulový mikrofون (za předpokladu, že všechny použité mikrofony jsou kulové a důležité informace v signálu jsou přibližně do 12 kHz, od stoupající frekvence se začnou mikrofony odečítat kvůli malé vlnové délce zvuku). Při zvolení této metody (ze

čtyř mikrofónů spočítáme jeden) je nutné zamyslet se nad tím, jestli odmocněním neztratíme podstatné informace o fázi apod.

Jednotlivé rovnice pro definování B-formát signálu potom vypadají následovně [11]:

$$\begin{aligned}x(t) &= \cos(\alpha) \cos(\beta) s(t), \\y(t) &= \sin(\alpha) \cos(\beta) s(t), \\z(t) &= \sin(\beta) s(t), \\w(t) &= \frac{1}{\sqrt{2}} s(t),\end{aligned}\tag{1.50}$$

kde  $\alpha$  je úhel v horizontální rovině a  $\beta$  je elevace zdroje zvuku oproti poli ve vertikální rovině,  $s(t)$  značí nahraný signál. Z obrázku je patrné, že ideálního B-formát signálu nikdy nedosáhneme, a to kvůli rozměrům mikrofónů.

### A-formát signálu

A formát signálu je složený ze čtyř signálů  $Lf, Lb, Rf, Rb$ , přičemž názvy jsou odvozeny pro jednotlivé směry, odkud by měl zvuk přicházet tzn.  $Lf$  je levo-přední (left-front-up), levo-dolní (left-bottom-down) je  $Lb$  a obdobně s  $Rf$  a  $Rb$  s tím, že jsou směřovány vpravo. Pokud je k dispozici nahraný signál A-formátu, potom se tento signál dá jednoduše přepočítat pomocí rovnic: [8]

$$\begin{aligned}W' &= Lf + Rf + Rb + Lb \\X' &= Lf + Rf - Rb - Lb \\Y' &= Lf - Rf - Rb + Lb \\Z' &= Lf - Rf + Rb - Lb\end{aligned}\tag{1.51}$$

Avšak pro správné nalezení vektorů  $W, X, Y, Z$  je nutné filtrovat výstup tohoto přepočtu, přičemž pro filtraci  $W$  filtr 0. řádu a pro  $X, Y, Z$  filtr 1. řádu.

### 1.3.10 Energická metoda s STFT

Hlavním předpokladem pro tuto metodu je, že směr umístění zvuku je v opačném směru než intenzita vektoru zvuku. Vzhledem k tomu, že energie zvuku je rozložená ve spektru a čase, je dobré spektrum zvuku rozdělit do frekvenčních úseků (anglicky „bin“). K tomu slouží STFT (Short Time Fourier Transform), která pomocí okna rozdělí signál a následně provede Fourierovu transformaci. Typy oken můžou

být různé, např. pravoúhlé, Kaiserovo nebo Hanningovo okno a další. Hlavním parametrem okna je jeho velikost. Pro občasné, nestacionární signály by mělo mít okno malou velikost, aby algoritmus pracoval správně. Kvůli nedokonalosti některých oken je nutné zadat ještě STFT overlapping (překryv), jinak hrozí ztráta dat u specifických signálů. [12] [13] Po provedení STFT můžeme matematicky zapsat vektor intenzity pro každý úsek takto:[11]

$$\begin{aligned} I_x(t, f) &= \frac{1}{\sqrt{2}z_0} \operatorname{Re} \{W^*(t, f) \cdot X(t, f)\} \\ I_y(t, f) &= \frac{1}{\sqrt{2}z_0} \operatorname{Re} \{W^*(t, f) \cdot Y(t, f)\} \\ I_z(t, f) &= \frac{1}{\sqrt{2}z_0} \operatorname{Re} \{W^*(t, f) \cdot Z(t, f)\}. \end{aligned} \quad (1.52)$$

V těchto rovnicích  $z_0$  je akustickou impedancí vzduchu,  $t$  je čas,  $f$  je frekvence daného úseku,  $(.)^*$  značí komplexně sdružené číslo,  $X(t, f)$ ,  $Y(t, f)$ ,  $Z(t, f)$  a  $W(t, f)$  jsou signály  $x(t)$ ,  $y(t)$ ,  $z(t)$  a  $w(t)$ , na které byla již aplikována Fourierova transformace. Jednotlivé vektory ukazují opačný směr přicházejícího zvuku v jednotlivých osách, takže je nutné je obrátit. Výsledný úhel v horizontální rovině potom můžeme získat pomocí: [1] [11]

$$\alpha(t, f) = \tan^{-1} \left[ \frac{-I_y(t, f)}{-I_x(t, f)} \right], \quad (1.53)$$

pro hodnoty  $I_y \geq 0$ . Pro hodnoty  $I_y < 0$  potom platí tato rovnice:[11]

$$\alpha(t, f) = \tan^{-1} \left[ \frac{-I_y(t, f)}{-I_x(t, f)} \right] - 180^\circ. \quad (1.54)$$

Pokud bychom chtěli získat elevaci (úhel ve vertikální rovině), potom použijeme rovnici:[11]

$$\beta(t, f) = \tan^{-1} \left[ \frac{-I_z(t, f)}{\sqrt{I_x(t, f)^2 + I_y(t, f)^2}} \right]. \quad (1.55)$$

Po spočítání úhlu každého frekvenčního úseku v čase, je nutné určit úhel celkový. Tento algoritmus předpokládá, že v každém frekvenčním úseku bude dominantní pouze jeden zdroj zvuku, což jistě klade velký důraz na dobrý odstup signálu od šumu (SNR). Dále může být problém se zvuky s podobným spektrem, jelikož mohou být reprezentovány jako jeden zdroj, zvláště, pokud se rozhodneme počítat pouze v určité části frekvenčního spektra pro urychlení a zjednodušení. Pokud budeme počítat s celým spektrem a jedním zdrojem, potom je nejčastěji zastoupený úhel ten hledaný. Při více zdrojích signálu toto však nemusí platit, avšak hledaný signál by měl být v některých frekvenčních úsecích nejčtetnější. Tato metoda se jeví jako dobrá pro



detekci pohybu signálu s předem známým frekvenčním spektrem. Matematický zápis pro hledání výsledného úhlu je následující: [1] [11]

$$N(\alpha) = \sum_{k=0}^K p(S(t, f_k)|\alpha), \quad (1.56)$$

kde  $K$  je počet frekvenčních úseků pro  $\alpha \in (-\pi, \pi)$ ,  $S$  je spektrum zvukového signálu,  $t$  značí index časového úseku,  $f$  je frekvenční úsek a  $p(S(t, f_k)|\alpha)$  je výraz pro nalezení pravděpodobnosti úhlu posunu  $\alpha$  přicházejícího signálu vůči poli. Po spočtení všech pravděpodobností úhlu vybereme celkový nejpravděpodobnější úhel  $\alpha_{\text{vysledna}}$  přicházejícího zvuku takto: [11]

$$\alpha_{\text{vysledna}} = \arg \max N(\alpha). \quad (1.57)$$

### 1.3.11 Metoda Beamforming

Metody hledání příchodu směru zvuku, patřící do rodiny beamformerů, jsou založeny na zlepšení sčítání chtěných signálů zatímco s ostatními zdroji signálů zachází jako s ruchem. Při použití s mikrofonním polem používají prostorovou filtraci, přičemž filtrace je prováděna na vstupním signálu s následnou kombinací výstupů (costructive combining–konstruktivní kombinace) a zahození (destructive combining–destruktivní kombinace) signálů z okolních ruchů podle jejich polohy pro získání požadovaného signálu. Beamforming může být využit jak v časové oblasti, tak frekvenční. V časové oblasti jsou využity FIR filtry na každý jednotlivý signál. Ve frekvenční oblasti je vstupní signál rozčleněn do jednotlivých frekvenčních pásem pomocí STFT a každé pásmo je filtrováno odděleně. V posledních letech byla na metodách beamformingu intenzivně zkoušena separace mluveného slova, a to díky mnoha užitím. [24]

#### Metoda Beamforming–DAS

Metoda DAS (Delay and Sum – „zpozdi a sečti“) je nejjednodušší z této rodiny. Je metodou deterministickou, filtry jsou nezávislé na vstupním signálu a mají pevnou výslednou odezvu (nepřízpůsobují se). Delay and sum (DAS) – jak napovídá název, v této metodě jsou vstupní signály zpožďovány a následně sčítány tak, aby přicházející signál z určitého směru byl koherentní. Předpokládá se, že případné rušení signálu se odečte právě destruktivní kombinací. Vstupní signál může být ve složitějších aplikacích rozložen na úzká frekvenční pásma a zpoždění na jednotlivých mikrofonech provedeno posunutím fáze signálu v každém pásmu. [24]

Za předpokladu, že počet mikrofónů  $M$  je rozestavený lineárně od sebe o rozdíl vzdálenosti  $\Delta_n$  a v řadě, potom na každém mikrofónu vznikne zpoždění a útlum.

Pokud je mikrofonní pole ve vzdáleném poli, je možné dopočítat čas dopadu vlny na mikrofony. [15][25]

$$y(k) = \sum_{n=0}^{N-1} w_n \cdot x_n(k - \tau_n) \quad (1.58)$$

je potom rovnicí, která popisuje vztah výstupního signálu  $y(k)$  a vzorků signálu  $x_n(k)$  z každého mikrofonu, který je zpožděn o  $\tau_n$  vzorků. Je možné potom takový signál před sčítáním váhovat díky proměnné  $w_n(k)$ . Pokud je vzdálenost zdroje zvuku od pole 5–10krát větší než velikost pole (závisí na charakteru signálu, potom je zpoždění na každém mikrofonu dáno rovnicí: [25]

$$\tau_n = \frac{\Delta_n \cdot \sin \theta}{c}, \quad (1.59)$$

kde  $\theta$  je úhel dopadu a  $c$  je rychlost zvuku a  $\Delta_n$  vzdálenost mezi referenčním mikrofonom a mikrofonom na pozici  $n$ .

Pro prostorové uskupení mikrofónů v kruhovém poli lze tyto rovnice potom vyjádřit jinak. Pokud pole o osmi mikrofonech rozdělíme v polovině na dvě části, úhel mezi každým mikrofonom je  $k2\pi/N$  vzhledem k příchodu zvuku  $\theta$ . Potom zpoždění mezi mikrofony můžeme vyjádřit jako: [26]

$$\Delta_k = \frac{r}{c} \cos \left( \theta - \frac{k2\pi}{N} \right), \quad (1.60)$$

kde  $r$  je poloměr kruhu pole,  $N$  je počet mikrofónů, přičemž  $k = 1, \dots, N$  značí index mikrofonu. Vztah mezi  $r$  a vzdáleností mezi mikrofony  $d$  je následující

$$d = 2r \sin \left( \frac{\pi}{N} \right), \quad (1.61)$$

Vzhledem k často nedostačnému výkonu této metody v reálném prostředí, byla vyvinuta metoda Filter and Sum (FAS).

### Metoda beamforming–Filter and Sum

Tato metoda je postavena na podobném principu jako Delay and Sum. FAS však nejdříve vstupní signál na každém mikrofonu filtruje FIR filtrem řádu  $M$  a potom je sečte. Takto získaný výstup je vhodnější pro více prostředí. [24]

Jednoduše jej můžeme zapsat v časové doméně jako konvoluci filtru  $w_m(l)$  a příchozího signálu  $x_m(n - \tau_n)$  jako: [27]

$$y(k) = \sum_{m=1}^M \sum_{l=0}^{L-1} w_m(l) \cdot x_m(n - \tau_n). \quad (1.62)$$

Při převedení do frekvenční oblasti a následujícím zápisu:

$W(\Omega) = [W_1(\Omega), W_1(\Omega), \dots, W_N(\Omega),]$  a  $X(\Omega) = [X_1(\Omega), X_2(\Omega), \dots, X_N(\Omega)]$  je výstup signálů z pole

$$Y(\Omega) = W(\Omega)^H \cdot X(\Omega). \quad (1.63)$$

### Beamformery postavené na statických údajích

Takové beamformery jsou založeny na statických údajích o vstupním signálu a jeho rušení. Přicházejícími daty filtr upravuje své funkce, tak aby byly optimální vůči chtěnému signálu. Funkce se mohou měnit pomocí více kritérií. Mohou to být například: maximální odstup signálu od šumu (MSNR–Maximum Signal to Noise Ratio), minimální střední kvadratická chyba (MMSE–Minimum Mean–Squared Error), odezva bez zkreslení s minimální odchylkou (MVDR–Minimum Variance Distortionless Response), lineární omezení s minimální odchylkou (LCMV–Linear Constraint Minimum Variance).

#### 1.3.12 Algoritmus ESPRIT

Algoritmus ESPRIT je možné aplikovat na mikrofoni pole, které obsahuje dvě identické a stejně orientované dílčí pole. Jedno dílčí pole je posunutá „replika“ toho druhého a tento vektor posunu musí být známý, nicméně geometrie každého pole může být neznámá.

Řídící vektor otáčení pro dvě dílčí pole může vyjádřen jako: [17]

$$a_1(\theta) = J_1 a(\theta), \quad a_2(\theta) = J_2 a(\theta), \quad (1.64)$$

kde  $J_1$  a  $J_2$  jsou matice, které mají v každém řádku jeden vstup je roven jedné a ostatní jsou rovny nule. Jak bylo řečeno výše, obě dvě dílčí pole jsou identické a stejně orientované. Posun jednoho pole potom provedeme kolem osy  $x$ . Potom řídící vektor otáčení obou polí může být zapsán v závislosti na druhém takto: [17]

$$a_1(\theta) = a_2(\theta) e^{j(2\pi/\lambda)d_x \sin(\theta)} \quad (1.65)$$

přitom  $d_x$  značí „odchylku“ mezi dvěma dílčími poli. Přepis těchto vztahů vypadá následovně:[17]

$$A_1(\theta) = J_1 A(\theta), \quad A_2(\theta) = J_2 A(\theta), \quad (1.66)$$

a zároveň

$$A_1(\theta) = A_2(\theta) Q(\theta), \quad (1.67)$$

kde

$$Q(\theta) \triangleq \text{diag} \left\{ e^{j(2\pi/\lambda)d_x \sin(\theta_1)}, \dots, e^{j(2\pi/\lambda)d_x \sin(\theta_L)} \right\}. \quad (1.68)$$

Směr přichozícího zvuku lze získat z diagonálních prvků  $Q(\theta)$ .

## 1.4 Matematické funkce

### 1.4.1 Diskrétní Fourierova transformace (DFT)

Důležitou funkcí pro signálové zpracování je DFT (Discrete Fourier Transform), která převádí signál z časové oblasti do frekvenčního spektra složek. Než je možné provést DFT, je nutné na vstup funkce poslat diskrétní signál, nemůžeme použít spojitou funkci. V praxi se používá vzorkování, které nám uloží hodnoty signálu  $s(t)$  v určitých diskrétních časových intervalech daných vzorkovací frekvencí.[1]

Navzorkovaný signál musí splňovat Nyquistův (Shannonův) teorém (1.69), který říká, že vzorkovací frekvence musí být více jak dvakrát větší než nejvyšší frekvence harmonických složek obsažených ve vzorkovaném signálu [7]

$$2f_{vz} > f_{\max}, \quad (1.69)$$

kde  $f_{vz}$  je vzorkovací frekvence a  $f_{\max}$  maximální frekvence harmonických složek. Pokud tento teorém není dodržen, potom dochází ke zkreslení signálu zvaný aliasing (překryvu kmitočtů vzorkovaného signálu), což je nežádoucí, často se proto používá antialiasingový filtr, který má za úkol odstranit kmitočty vyšší než  $\frac{f_{vz}}{2}$ . Příkladem může být nahrávání signálu, který chceme korelovat. Kdybychom podcenili vzorkovací frekvenci, mohlo by dojít k zašumění signálu a šum by byl navíc závislý, což by značně ovlivnilo výsledek, jak popsáno v 1.4.3. Aliasing se vyskytuje ve více oborech, v hudebním zpracování, filmu nebo třeba v grafice.

Diskrétní Fourierova transformace pracuje s konečným signálem o  $N$  vzorcích v diskrétním čase, narozdíl od spojitě Fourierovy transformace. Proto je nutné přepsat integrál ve spojitém čase (1.70) [7]

$$s(t) = \int_{-\infty}^{\infty} \hat{s}(f) \cdot e^{i2\pi ft} dt, \quad (1.70)$$

na sumu diskrétních vzorků

$$\hat{s}_n = \sum_{k=0}^{N-1} s_n \cdot e^{-i2\pi ik \frac{n}{N}} = \sum_{k=0}^{N-1} s_n \cdot W_N^{kn}. \quad (1.71)$$

### 1.4.2 Krátkodobá Fourierova transformace (STFT)

Krátkodobá Fourierova transformace (Short-Time Fourier Transform) je algoritmus, který vychází FFT (Fast Fourier Transform), akorát zpracovává signál po částech, které jsou vymezeny tzv. oknem. [1]

## STFT v diskrétním čase

Vstupním signálem je již navzorkovaný signál v čase, což provádí A/D převodník. Každý vzorek v čase je rozdělen na několik pásem pomocí okna a je převeden do frekvenční oblasti pomocí FFT. Matematicky lze STFT napsat jako:[14]

$$\text{STFT} \{x[n]\}(m, \omega) \equiv X(m, \omega) = \sum_{n=-\infty}^{\infty} x[n] \cdot w[n-m] \cdot e^{-i\omega n}, \quad (1.72)$$

kde  $x[n]$  je vstupní signál a  $w[n]$  je okno, kterým se signál násobí. Výraz  $e^{-i\omega n}$  značí Fourierovu transformaci, v praxi se potom často zavádí rychlá Fourierova transformace. Pokud je potřeba zmenšit okno  $w$  k dosažení lepších výsledků, může se posunout o parametr  $m$ , kterým by měla být nespojitá hodnota. [14]

## Vybraná okna pro STFT

Okna pro STFT mohou být různá, např. obdélníkové, Bartlettovo, Welchovo, Hanningovo, Hammingovo, Blackmann–Harrisovo, či Nuttallovo. Podrobný přehled a zhodnocení je k nalezení v referenčním zdroji [13]. Já se zmíním pouze o pár oknech, které jsou vhodné pro algoritmus energetické metody. Jen nutné si uvědomit, že čím složitější okno, tím větší bude nárok na výkon počítače a také je nutné znát, k jakému účelu okna používáme, jestli špatně zvolené okno může ovlivnit výsledky měření.

### Hanningovo okno

Hanningovo okno je jedno z oken, která jsou založena na váhovacím součtu konstantního výrazu a kosinovou funkcí výrazu. Zapsat jej můžeme jako [13]

$$z = \frac{2\pi j}{N}, \quad j = 0 \dots N-1, \quad (1.73)$$

$$w_j = \frac{1 - \cos(z)}{2} = \cos^2\left(\frac{z - \pi}{2}\right). \quad (1.74)$$

Hanningovo okno se používá v mnoha případech, kde nevádí zkreslení spektra sinusové vlny (například u měření šumu).

### Blackman–Harrisovo okno

Základ algoritmu je ve sčítání kosinové funkce. Existuje více druhů tohoto okna, proto popíši jeden vzorový příklad. Jde o takzvané „92 dB Blackman–Harrisovo okno“, které bylo nakonfigurováno speciálně tak, aby boční laloky v přenosové funkci, které sousedí s hlavním, byly malé.[13]

$$z = \frac{2\pi j}{N}, \quad j = 0 \dots N - 1, \quad (1.75)$$

$$w_j = 0,35875 - 0,48829 \cos(z) + 0,14128 \cos(2z) - 0,01168 \cos(3z), \quad (1.76)$$

takto vypadá matematický zápis. Díky malým postranním lalokům se hodí i na detekci malých sinusových signálů, které jsou kmitočtově blízké velkým signálům. Například pokud Fourierovou transformací dostaneme velkou amplitudu základní frekvence a k tomu malé amplitudy dalších složených frekvencí. Další výhodou může být, že funkce má velký dynamický rozsah.

### 1.4.3 Vzájemná korelace

Hlavní matematickou funkcí pro určení směru je vzájemná korelace, která udává podobnost signálu vůči druhému (speciální případ je autokorelace, podobnost signálu sobě samému). Vstupem jsou dvě funkce, výstupem je jedna. Matematická definice je [1][6]

$$R_{xy}(\tau) = \int_{-\infty}^{\infty} x(t)y(t + \tau)dt, \quad (1.77)$$

Maximum této funkce je v bodě, kdy jsou si vstupní signály  $x(t)$  a  $y(t + \tau)$  nejvíce podobné, v nejlepším případě autokorelace stejné. Minima bude rovnice (1.77) dosahovat tehdy, když vstupní funkce budou mít shodné tvary, ale jedna z nich bude kladná a druhá záporná (budou v protifázi). Potom se výsledky budou blížit nule.

Z popisu výše je patrné, že přidaný šum do signálu musí být náhodný a nezávislý na vstupním signálu.[5] Také právě proto je výhodné převádět vstup do frekvenční roviny Fourierovou transformací, neboť potom můžeme přímo porovnávat jednotlivé složky signálu a ne celý, těžko bychom porovnávaly dva signály v časové doméně.

## 2 VYHODNOCENÍ PRAKTICKÉ ČÁSTI

V této sekci budu popisovat praktické uskutečnění teoretických poznatků výše. Jako programovací prostředí byl použit simulační program Matlab, snímání nahraných zvuků bylo realizované v bezodrazové komoře nebo byly uměle vytvořeny. K veškerým měřením bylo použito mikrofونů AUDIX TM1, což jsou testovací a měřicí kulové mikrofony, které jsou dostupně v laboratoři FEKT. [1]

### 2.1 Popis měření a zpracování výsledků

Nahrávací vzorkovací kmitočet jsem zvolil 48 kHz, přičemž je jasné, že čím větší vzorkovací kmitočet bude, tím bude i přesnější měření. Na druhou stranu nutno podotknout, že při vyšším vzorkování bude i větší velikost souboru a větší náročnost na zpracování dat. Můžeme si totiž lehce odvodit vzorec, jaký může být největší časové zpoždění signálu vyjádřené v počtu vzorků  $\Delta s_n$  mezi dvěma mikrofony na vzdálenosti  $d$ : [1]

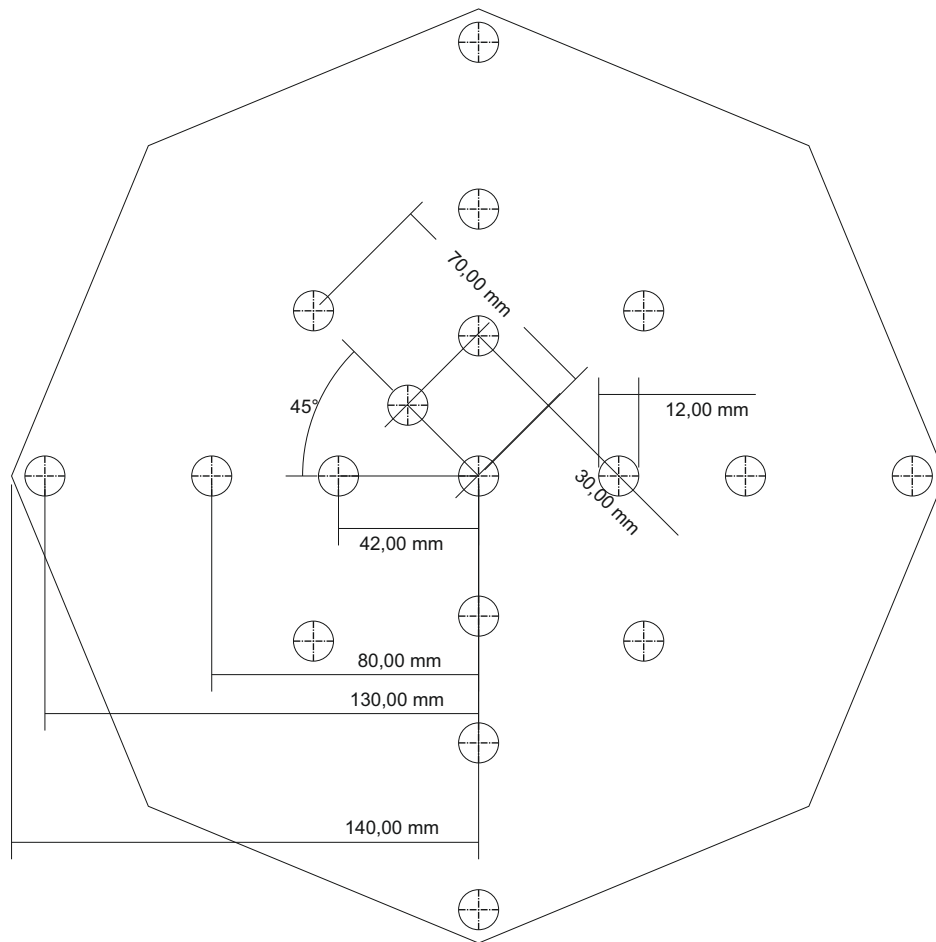
$$\Delta s_n \doteq \frac{f_s d}{c}, \quad (2.1)$$

kde  $f_s$  je vzorkovací frekvence,  $d$  vzdálenost mezi mikrofony a  $c$  je rychlost zvuku. Z tohoto vztahu plyne, že na větší vzdálenost dokážeme zaznamenat větší zpoždění, samozřejmě musíme stále předpokládat, že mikrofony jsou ve volném poli. Je také nutné správně určit rychlost zvuku v daném prostředí, protože na vzdálenosti 5 cm může mikrofون zaznamenat (při vzorkovací frekvenci 48 kHz) 6 vzorků při rychlosti zvuku  $345 \text{ m} \cdot \text{s}^{-1}$  a 7 vzorků při rychlosti  $340 \text{ m} \cdot \text{s}^{-1}$ . A na počtu vzorků je přímo úměrné dělení úhlů, tzn. rozmezí  $0^\circ$ – $90^\circ$  bude při 7 vzorcích rozděleno na 7 částí, které ale nebudou stejně velké (kvůli funkci  $\arccos$ ).

Pro hledání směru zvuku v malém mikrofonním poli jsem zvolil křížové pole, dále kruhové pole o průměru 20 cm a také pole čtyř mikrofونů pro B-formát signálu. Zvuk jsem nahrával v bezodrazové místnosti s dozvukem  $RT60 < 0,3 \text{ s}$  v budově FEKT, proběhlo i měření ve venkovních prostorách. Jako pokusný signál jsem volil bouchnutí dveří, tříštění skla a výstřel z pistole, a to opět kvůli jejich frekvenčnímu spektru. Pohyb zvuku byl simulován pomocí natáčení pole pomocí točny, který je dostupný v laboratoři FEKT. Točna je spojen s dálkovým ovládáním, jde na něm nastavit přesný úhel, o kolik se mikrofonní pole posune. Z předem vypracovaných výpočtů jsem se rozhodl rotovat pole v intervalu  $\langle 0^\circ, 90^\circ \rangle$ , protože v intervalu úhlů  $(90^\circ, 360^\circ)$  by měl být signál stejný, ale s jinými znaménky. [1]

### 2.1.1 Parametry měření s vyrobeným mikrofonním polem

#### Návrh a realizace držáku mikrofonního pole



Obr. 2.1: Původní navržený držák mikrofonního pole [1]

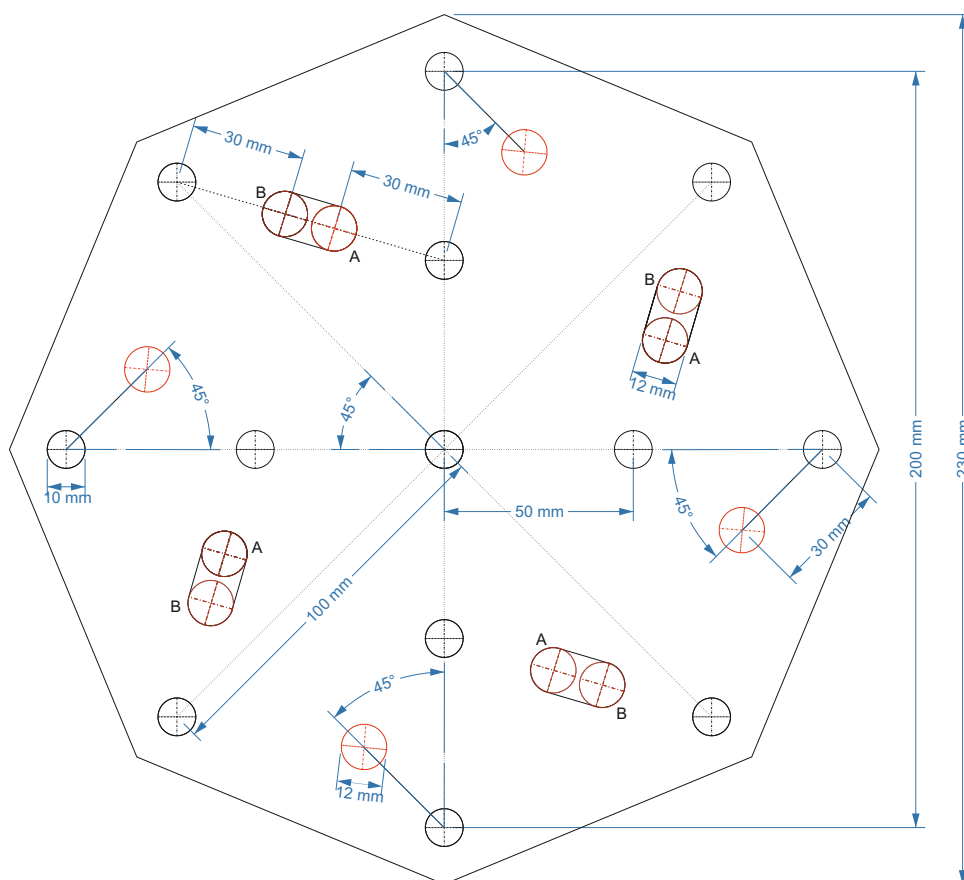
Abych zlepšil výsledky měření, a zároveň zjednodušil instalaci mikrofonů, tak jsem navrhl držák mikrofonního pole. Lze na něj instalovat hned několik typů mikrofonních polí, křížové, kruhové a pole pro měření energetickou metodou. Otvor uprostřed slouží k zachycení na stojan, ostatní pro mikrofony. Jednotlivé parametry (vzdálenosti děr) jsou změřeny a navrženy speciálně pro měřicí mikrofony, průměr děr je navržen pro klasické držáky. Křížové pole je popsáno již na obr. 2.3. [1]

Kruhové pole má průměr 20 cm. Při energetické metodě jsou mikrofony uchycovány tak, aby byly co nejblíže u sebe. Původní návrh držáku byl kruhový a z plechu. Nakonec pro jednodušší výrobu jsem volil osmiúhelník a materiálem je plast o tloušťce 3,6 mm. Držák na pole byl vyroben a ozkoušen. Pro zlepšení bych nechal více prostoru pro kabely mikrofonu, protože u tohoto návrhu je nutné použít zalomené XLR



kabely. Co je nutné dále řešit je prokluzování držáků samotných mikrofonů, nicméně stačí pro řešení tohoto problému použít podložky. Jinak mikrofony jsou ve správných vzdálenostech i výšce, takže hlavní účel držák plní bez problémů, snad u kruhového pole se špatně odhaduje poloha mikrofonů ležících v úhlu  $45^\circ$  od středu, ale zde pomáhá zvolený tvar osmiúhelníku držáku. [1]

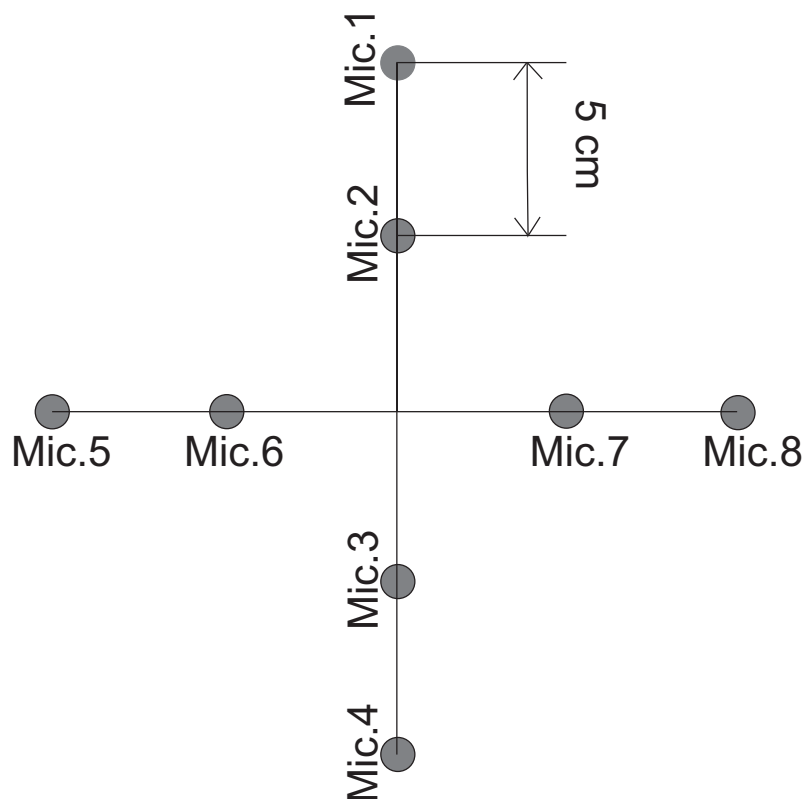
### Návrh na vylepšení držáku mikrofonního pole



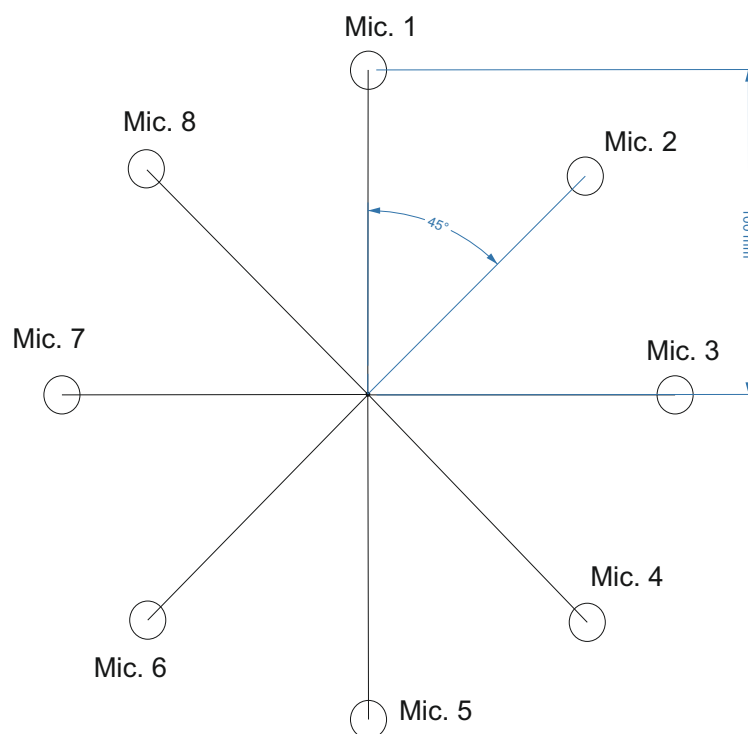
Obr. 2.2: Vylepšené navržený držák mikrofonního pole

V diplomové práci bylo dílčím cílem odstranit problémy předchozího držáku a to hlavně v těchto aspektech: nemít nutnost zalomeného kabelu XLR, dále omezit prokluzování držáků mikrofonů, zjednodušit přechody mezi jednotlivými typy polí a hlavně lépe definovat a zamezit geometrickým nepřesnostem jednotlivých typů polí. Proto jsem navrhl držák, jehož popis je na obr. 2.2. Tento držák umožňuje dva typy rozestavení mikrofonních polí, křížové a kruhové. Oproti původnímu držáku

byla vynechána možnost pole pro B-formát signálu, jelikož přepočít ze čtyř mikrofonů kulového typu není vhodný (výsledky energetické metody jsou níže). Objímky na mikrofony se montují do otvorů o průměru 12 mm, přičemž hlavy mikrofonů se prostrkují zespodu do děr o velikosti 10 mm. Tato metoda byla zvolena kvůli možnosti mít jakýkoli XLR kabel a hlavně kvůli přesnosti geometrie polí, nicméně nevýhodou je, že se více promítne odraz zvuku od plastové destičky. Proto je vhodné tento odraz utlumit tlumícím materiálem, který se nalepí na držák (jedná se o utlumení signálu od kmitočtu 2 kHz a výš). Každopádně toto tlumení má spíše menšinový vliv na přesnost, protože signál z mikrofonů je ořezáván filtrem dolní propustí prvního řádu o mezním kmitočtu 4 kHz. Na obrázku jsou naznačeny možnosti dvou poloh mikrofonu (A,B). Pokud je mikrofon na pozici A, lze jej použít jako součást křížového pole, v opačném případě (poloha B) je to pole kruhové (popřípadě je možné použít dva mikrofony).



Obr. 2.3: Schéma rozestavení mikrofonů v křížovém poli. [1]



Obr. 2.4: Schéma rozestavení mikrofonů v kruhovém poli. [1]

### 2.1.2 Zpracování v Matlabu

V této části se zabývám postupem, který jsem naprogramoval. Nejdříve je nutné nahrané vzorky nahrát do proměnných. Jelikož se jedná již o diskretní signál, tak Matlab vytvoří matici o  $1 \times N$  polí, kde  $N$  je počet samplů. Pro další zpracování je tento formát nevhodný, tudíž je lepší z něj udělat matici o  $N \times 1$  polích, což se provede inverzí matice.

```
[y,fs] = audioread ();
Y=[y, y1,y2,y3,y4,y5,y6,y7];
Y=Y';
y= Y(1,:);
```

Audioread je příkaz k načtení zvukového souboru do proměnné  $y$  se vzorkovací frekvencí  $fs$ . Abych mohl jednoduše zpracovat výsledky z celého pole najednou, vytvořil jsem si jednoduchou funkci, která sama spočítá pohyb předmětu. [1]

```
function pohyb = ALL_methods(X, Y, pole, metoda, fs, c)
```

Matice se zvukem musí být již zpracované tak, aby vstupní matice měla v řádcích jednotlivé mikrofony v poli. Vstupy jsou  $X$ ,  $Y$ . Jsou zde zahrnuty metody GCC

(Generalized Cross-Correlation), PHAT (Phase Transform) a ML (Maximum Likelihood).  $X$  je první poloha předmětu (nahraný zvuk mikrofonním polem) a  $Y$  je druhá poloha, v matici může být až osm mikrofonů. Volby: pole = 0 je křížové pole, pole = 1 je kruhové pole; metoda = 0 je GCC, metoda = 1 je PHAT, metoda = 2 je ML. Proměnná  $fs$  je vzorkovací frekvence v Hz a  $c$  je rychlost zvuku v m/s.

## Preprocessing

K získání přesného výsledku jsem naprogramoval skript, který jednotlivé signály z mikrofonů upravuje tak, aby náročnost algoritmu byla co nejmenší, zároveň však aby nebyla zanedbána přesnost. Zkoumání přesnosti na malém počtu vzorků je velmi složité na interpretaci. Je nutné si uvědomit, že pokud délka signálu je 1000 vzorků při samplovací frekvenci 48 kHz, tak se jedná o čas 0,021 s. Pokud tedy je předmětem zkoumání pohyb, který změnil svoji polohu o 90° vůči poli za tak krátký čas, je otázkou, jestli takový pohyb v reálné situaci může nastat a pokud, tak jak často. Nicméně takový pokus může sloužit jako otestování hraničních situací.

Abych zamezil výběru části nahraného souboru, kde není signál dostatečně silný, vzal jsem oblast, kdy amplituda signálu nabývá maxima.

```
index_maxima = find(x1 == max(x1),1) - 1;

X = X(:,index_maxima-oblast_maxima:...
        index_maxima+oblast_maxima-1);

Y = Y(:,index_maxima-oblast_maxima:...
        index_maxima+oblast_maxima-1);
```

Počet vzorků kolem maxima je možné ovlivnit proměnnou `oblast_maxima` (je dobré volit takovou oblast, která bude násobkem odmocniny ze dvou, kvůli FFT). Dále není nutné zpracovávat signál v celém spektru (jednotlivé podmínky závisí na metodách). Signál jsem tedy omezil horní a dolní propustí (použit byl fázovací článek) naprogramovanou funkcí, která projde všechny mikrofony:

```
function Y = preprocessing_m(Y,p_kanalů,fs),
```

kde  $Y$  je matice se vstupním signálem,  $p\_kanalů$  je počet kanálů a  $fs$  je vzorkovací frekvence.

## Metody GCC, PHAT a ML

Nyní můžeme přistoupit k samotnému zjišťování časového zpoždění. Jádrem metody metody GCC je převod signálu do frekvenční oblasti (aby se počítala FFT správně,

musí být délka signálu násobkem mocniny dvou, pokud funkce je postavena na Cooley–Tukey algoritmu, což v Matlabu je).

```
X1 = fft (x);
X2 = fft (x1);
Sx1 = (X1).*conj(X2);
%výsledná korelace v počtu vzorků
fx = ifft (Sx1);
```

Funkce FFT, IFFT (zpětná rychlá Fourierova transformace) jsou v Matlabu již naprogramované, a proto jsem jich využil. Pro správný výstup křížové korelace musí mít vstupní signál stejnou délku, což je třeba ošetřit, například stejnou časovou délkou nahraných vzorků nebo uměle zarovnat matici. Následně už můžeme provést FFT obou signálů ze vstupu a vyjádříme z nich výkonovou spektrální hustotu, kterou potom vydělíme její velikostí (absolutní hodnotou). Když takto vydělenou výkonovou spektrální hustotu převedeme zpět do časové osy pomocí IFFT, dostaneme korelační funkci, jejíž maximum značí časovou prodlevu mezi jednotlivými signály na vstupu.

Jak bylo zmíněno výše, algoritmus PHAT vlastně jenom jiným způsobem upravuje tvar výsledného spektra

```
X1 = fft (x);
X2 = fft (x1);
Sx1 = (X1).*conj(X2);
Sx2 = abs(Sx1);
Fx = Sx1./Sx2;
fx = ifft (Fx);
%zpětná transformace , následuje hledání maxima
```

Pro funkci ML vypadá kód funkce takto:

```
X1 = fft (x);
X2 = fft (x1);
Gx12 = (X1).*conj(X2); %různá výkonová spektra
Gx11 = (X1).*conj(X1);
Gx22=(X2).*conj(X2);
delta2 = ((Gx12).^2)/(Gx11.*Gx22); %výkonové spektrum
Fiml= (1./abs(Gx12)).*(abs(delta2)./(1-abs(delta2)));
%výsledný filtr
fx = ifft (Fiml.*X1.*conj(X2));
%zpětná transformace , následuje hledání maxima
```

Jedná se o přepis rovnic z kapitoly 1.3.5, a to (1.34), (1.35) a (1.36). Pro zjednodušení ovládání jsem vytvořil funkci, která provádí výše popsany algoritmus, ve tvaru:

```
function posun = faze(x,x1,d,fs),
```

kde  $x$  je signál z jednoho mikrofону,  $x1$  signál druhého mikrofону,  $d$  je vzdálenost mikrofónů v metrech,  $fs$  je vzorkovací frekvence. Pro různé metody jsou vstupní parametry stejné  $x, x1, d, fs$ , takže jsem funkce jednotlivých metod jenom jinak pojmenoval. Celý kód funkce metody GCC je v příloze A.1 a jednotlivé další metody v přílohách A.2 a A.3.

Při zjišťování posunu signálu musíme dát pozor, aby platila podmínka

$$\Delta t < \frac{d}{c}, \quad (2.2)$$

kde  $\Delta t$  je časové zpoždění,  $d$  vzdálenost mezi mikrofony,  $c$  značí rychlost zvuku. Je to nutné kvůli tomu, že funkce  $\arccos$  platí v reálných číslech pouze na intervalu  $\langle -1, 1 \rangle$ . Pokud tuto podmínku nedodržíme, Matlab vypíše výsledek v komplexním čísle, což není žádoucí, protože nedokážeme vypočítat zpoždění a dále je tato hodnota ignorována.

## Energetická metoda

Jak je již zmíněno v kapitole o Energetické metodě 1.3.10, tato metoda je založená na B-formátu signálu a STFT. Hlavička<sup>1</sup> mé naprogramované funkce je nadefinovaná takto:

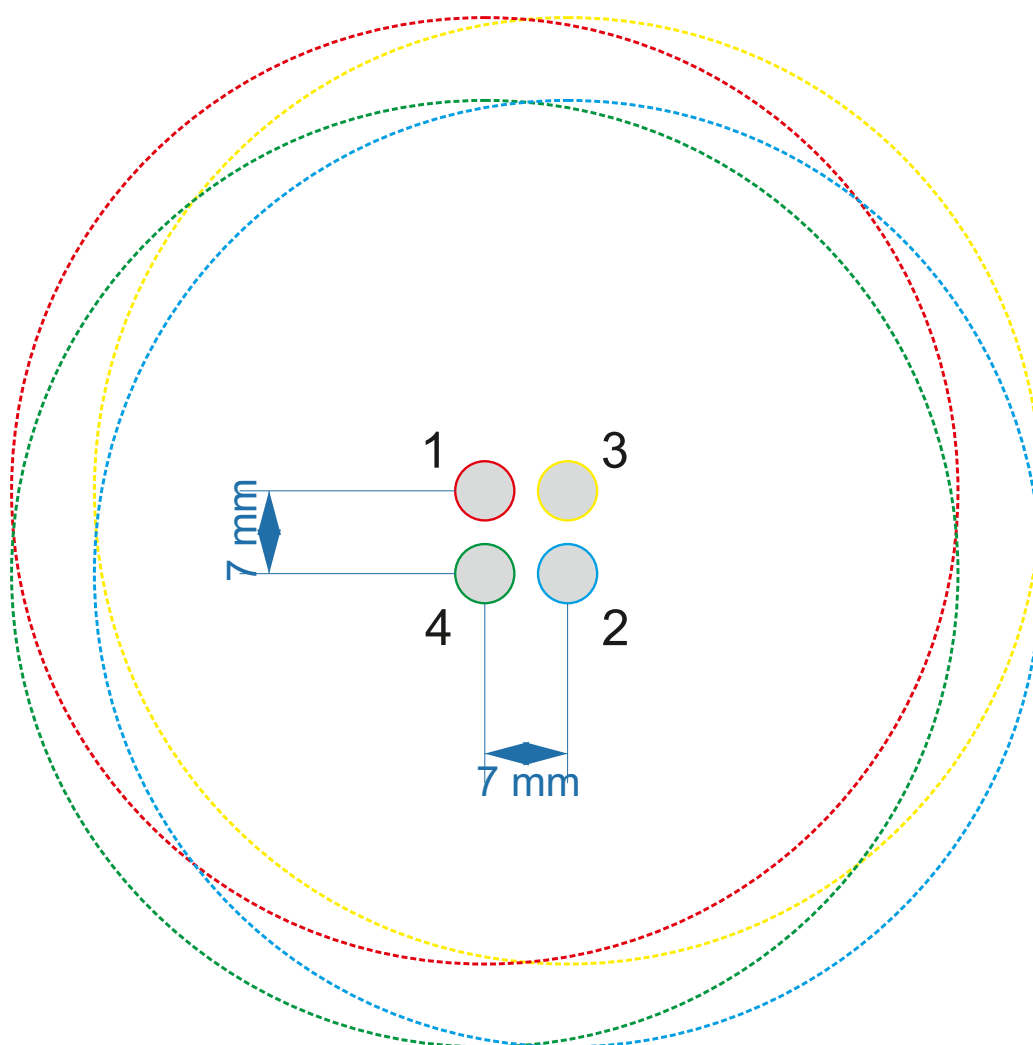
```
function energetick = E_m (x, posun, nfft, okno, fs)
```

Vstupní parametry jsou  $x$  (matice signálů 4 mikrofónů),  $posun$  (zadáva se ve vzorcích, definuje, o kolik se posune STFT v čase),  $nfft$  (počet binů FFT),  $okno$  (udává velikost okna při STFT) a  $fs$  (vzorkovací frekvence nahraného signálu). Matice vstupních signálů není ve správném formátu, proto ji upravit. Přepočítání signálu může být zapsán jako:

```
Xfb= x1+x2;
Ylr = x4+x3;
Wc = ((Xfb.^2)+(Ylr.^2)).^(1/2);
```

---

<sup>1</sup>Ve skriptech je tato metoda nazvaná Energeticka\_metoda místo  $E\_m$  (zde je název krácen kvůli formátování).



Obr. 2.5: Ukázka nastavení mikrofonů pro B-formát. Všechny mikrofony jsou kulového charakteru.[1]

Proměnné  $x_1$  a  $x_2$  je signál protilehlých mikrofonů, stejně jako  $x_3$  a  $x_4$  viz obr. 2.5. Vzhledem k chybějící možnosti získat přímo nahrávky v B-formátu, bylo vytvořeno pokusné pole o čtyřech kulových mikrofonech ve čtvercovém poli s co nejmenší délkou, které se následně muselo přepočítat na příslušný formát signálu. Nicméně tento přepočet se neosvědčil, výsledky nejsou přesvědčivé. Proto bylo nutné nahrát A-formát signálu a následně jej přepočítat podle následujících rovnic.

```
%B formát signálu
```

```
Lb=x(1,:);
```

```
Lf=x(2,:);
```

```
Rb=x(3,:);
```

```
Rf=(x(4,:));
```

```
Xfb = Lf+Rf-Rb-Lb;
```

```
Ylr = Lf-Rf-Rb+Lb;
```

```
Wc = Lf+Rf+Rb+Lb;
```

Po získání potřebného formátu signálu je nutné na něj aplikovat STFT. V kódu Matlabu může tato funkce vypadat například:

```
% perform STFT
```

```
for i = 1:pocetsloupce
```

```
%signál x od časového indexu 1 po časový index + okno,
```

```
xw = Xfb(indx+1:indx+okno);
```

```
windowf = xw.*blackhar;
```

```
% FFT
```

```
X = fft(windowf, nfft);
```

```
%uložení do matice výsledků
```

```
Xvysl(:, i) = X(1:pocetrad);
```

```
%posun indexu
```

```
indx = indx + posun;
```

```
end
```

Jedná se o jednoduchý cyklus od 1 do konce signálu, kdy se vezme část signálu, vynásobí se funkcí okna (v mém případě bylo použito Blackman–Harrisovo viz kapitola 1.4.2) a následně se provede Fourierova transformace. Potom už jen zbývá zapsat přepočítaný signál do matice výsledků a posunout se na další část signálu.

Předposledním krokem je vypočítat úhel, na které jednotlivé úseky (biny) STFT signálu ukazují. Podle rovnic (1.52) a (1.53) je možné nalézt je takto:



```

[q,w]=size(Wvysl);

for high = 1:w
Ix(:,high)=(1/(power(2,0.5)*z0))*real...
    ... (Xvysl(:,high).*conj(Wvysl(:,high)));

Iy(:,high)=(1/(power(2,0.5)*z0))*real...
    ... (Yvysl(:,high).*conj(Wvysl(:,high)));
end
[m,n]=size(Ix);
Vyslalfa = zeros(m,n);

for i = 1:n
    for j = 1:m
        if Iy(j,i)>=0
Vyslalfa(j,i)=atan((-Iy(j,i)./Ix(j,i)))*180/pi;
        else
Vyslalfa(j,i)=(atan((-Iy(j,i)./Ix(j,i)))-pi)*180/pi;
        end
    end
end
end

```

Podle teorie lze i určit úhel elevace zvuku, nicméně tento parametr není v předmětu této práce, proto jej do kódu nezahrnuji. Poslední částí této metody je vybrat z výsledné matice četnosti úhlů, kam ukazují jednotlivé biny. V kódu je tento problém řešen vektorem úhlů, který se nabývá hodnot od -360 do 360 (lineárně), přičemž stačí zjistit pouhým porovnáním, na jakém indexu se prvek výsledné matice shoduje s vektorem úhlů.

```

vec = -360:1:360;
[~,krok] = size(vec);
vys = vys - mod(vys,krok_fr)+1;
cetnost=zeros(floor(vys/krok_fr),krok);

```

Nejdřív je nutné zajistit, aby délka analyzovaného signálu byla násobkem okna, které jsem naprogramoval pro zrychlení algoritmu (rozdělení frekvenčních binů do větších částí). Využívám k tomu zbytek po dělení. Následně je řešena podmínka, která vymezuje interval úhlu, který byl nalezen (porovnává hodnotu daného prvku a vektoru úhlů). Zbývá už jen provést přepočty nejčetnějších indexů na stupně (index 1 je -360 stupňů) a vložit je do matice. A jelikož je možné, že četnost úhlů se může shodovat, výsledný formát matice úhlu je cell.

```

for time = 1:n
for p = 1:(vys-1)
    Vyslalfap = Vyslalfa(p,time)';
    for k = 1:krok-1;
        akprvek = floor(Vyslalfap);
        if (akprvek >= vec(k) && akprvek <= vec(k+1))
            cetnost(p,k) = cetnost(p,k)+1;
            break
        end
    end
end
end
energetic=cell(p,krok);
for i = 1: krok
    for k = 1:p
        cetnostp= cetnost(k,:);
        [~,energeticp_index]=find(cetnostp==max((cetnostp),[],2));
        %vyhledává jednotlivé četnosti úhlů
        energetic{k,i} = (energeticp_index(:)-720)/2;
    end
end
end

```

Vzhledem k tomu, že má výsledná matice velmi mnoho prvků, zvolil jsem metodu průměrování. Jediným problémem udělat průměr z hodnot je formát matice cell a to, že některé buňky obsahují více než jednu hodnotu (často „mylnou“). Postupoval jsem následovně: buňky s více prvky jsem zprůměroval a vytvořil průměr i s takto „chybnými“ údaji, které jsem následně vyloučil podmínkou, že pokud nejsou v rozmezí  $\pm 30^\circ$  od „chybného“ průměru, nezapočítám je.

```

for s = 1:l
    hranice = abs(floor(uhel));
    akvzorek =abs(prum_celkovy(s-cetnost,1));

    if akvzorek>=(hranice + 30) | akvzorek<=(hranice - 30)
        prum_celkovy(s-cetnost,:) = [ ];
        cetnost = cetnost+1;
    end
end
end

```

## 2.2 Výsledky

### 2.2.1 Výsledky zkušebního signálu

V této kapitole jsou shrnuty výsledky měření popsaného výše. Zpracování signálu probíhalo jak je popsáno v 2.1.2. V tab. 2.1 jsou výsledky časového posunu ( $\Delta t$ ) a odpovídající úhel ( $\theta$ ) dopadajícího naprogramovaného šumu, jehož korelaci je možno vidět na obr. 2.6. Maximum funkce je velice dobře zřetelné (obr. 2.6), protože se jedná o nasimulovaný nerušený signál v bezodrazovém prostředí. Počítal jsem výsledky pro vzdálenost mikrofónů  $d = 5, 10$  cm a  $f_s = 48,96$  kHz. Z tabulky lze vyčíst, že pro  $d = 5$  cm čím větší časový rozdíl v rozmezí  $\langle 0; 1,4583 \cdot 10^{-4} \rangle$  s, tím víc nelineárně roste úhel, což je dáno funkcí  $\arccos$ . Průměrná změna úhlu v rozmezí zpoždění od 0 do 5 vzorků cca.  $9^\circ$  a v oblasti od 5 do 7 vzorků  $18,752^\circ$ . V praxi to znamená, že je nutné snímat více mikrofóny než dvěma, aby byla chyba u větších časových rozmezí co nejmenší. [1]

Tab. 2.1: Posun úhlu v závislosti na zpoždění při vzorkovacím kmitočtu 48 kHz. [1]

Index maxima	$\theta$ [ $^\circ$ ]	$\Delta t$ [s]
0	90	0
1	98,1443	$2,0833 \cdot 10^{-5}$
2	106,4592	$4,1667 \cdot 10^{-5}$
3	115,1507	$6,25 \cdot 10^{-5}$
4	124,5181	$8,333 \cdot 10^{-5}$
5	135,0995	$1,0417 \cdot 10^{-4}$
6	148,2117	$1,25 \cdot 10^{-4}$
7	172,598	$1,4583 \cdot 10^{-4}$
8	180-29,268i	$1,666 \cdot 10^{-4}$

Tab. 2.2 popisuje časová zpoždění téhož šumu, akorát snímány dvojnásobnou vzorkovací frekvencí.

Jde vidět, že při dvojnásobné vzorkovací frekvenci bude přesnost určení úhlu dvakrát větší, jelikož se měřený interval rozdělí na dvakrát více vzorků. Na druhou stranu, čím vyšší kmitočet zvolíme, tím větší nároky klademe na parametry nahrávacího přístroje.

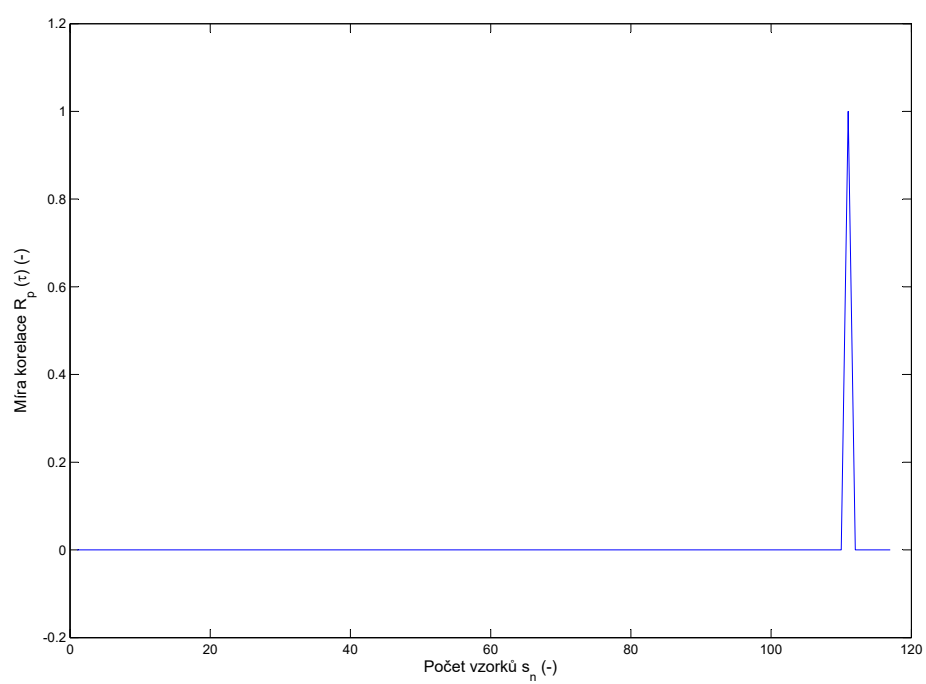
Co se týče vzdálenosti mikrofónů, tak z tabulek je možno vyčíst, že má na výsledek podobný vliv jako vzorkovací kmitočet, tzn. pokud vezmeme 5 cm rozestup, dostaneme maximálně 7 vzorků. Při 10 cm rozestupu je to 14 vzorků atd. Opět se

nám dvakrát zlepši výsledky měření. Je však nutné brát v potaz i praktickou stránku jako je sestavení pole, jestli se bude jednat o pevnou instalaci nebo přenosnou. [1]

Tab. 2.2: Posun úhlu v závislosti na zpoždění při vzorkovacím kmitočtu 96 kHz. [1]

Index maxima	$\theta$ [°]	$\Delta t$ [s]
0	90	0
1	94,0619	$1,04 \cdot 10^{-5}$
2	98,1443	$2,0833 \cdot 10^{-5}$
3	102,2689	$3,14 \cdot 10^{-5}$
4	106,4592	$4,1667 \cdot 10^{-5}$
5	110,7424	$5,2083 \cdot 10^{-5}$
6	115,1507	$6,25 \cdot 10^{-5}$
7	119,7247	$7,2917 \cdot 10^{-5}$
8	124,5181	$8,333 \cdot 10^{-5}$
9	129,6057	$9,375 \cdot 10^{-5}$
10	135,0995	$1,0417 \cdot 10^{-4}$
11	141,1843	$1,1458 \cdot 10^{-4}$
12	148,2117	$1,25 \cdot 10^{-4}$
10	157,0482	$1,3542 \cdot 10^{-4}$
14	172,598	$1,4583 \cdot 10^{-4}$
15	180-20,1531i	$1,5625 \cdot 10^{-4}$

Tyto dvě části měření prakticky dotvrzují vzorec (2.2). Pro měření z této simulace plyne, že použiji spíše pole kruhové, s referenčním mikrofonom uprostřed. Dále z tabulek vyplývá, že je nutné před měřením „zkalibrovat“ mikrofonní pole tak, abych dokázal správně určit úhel. Provedení bude takové, že do místa, které určím jako  $0^\circ$ , postavím zdroj zvuku a potom všechny mikrofony „vynuluji“ tím, že k nim přičtu nebo odečtu právě naměřený úhel.[1]



Obr. 2.6: Graf míry korelace zkušebního signálu (ukázka)[1]

## 2.2.2 Výsledky z měření v laboratoři s reálným signálem

Popis průběhu měření je v kap. 2.1. Pro hledání směru zvuku v malém mikrofonním poli jsem zvolil křížové a kruhové pole o osmi mikrofonech, dále potom B-formát mikrofونů pro energetickou metodu. Jednotlivé nákresy polí je vidět z obr. 2.3 pro křížové pole, pro kruhové pole 2.4, pro B-formát 1.1. Zvuk jsem nahrával v bezodrazové místnosti s dozvukem  $RT60 < 0,3$  s, která se nachází na půdě FEKT. [1]

Jako signály jsem použil nahrávku bouchnutí dveří, tříštění skla a výstřel z pistole kvůli jejich frekvenčním předpokladům (obsahují nízké a vysoké kmitočty v nahrávce). Testovací signály byly nahrány se vzorkovacím kmitočtem 44,1 kHz. Pohyb zvuku jsem simuloval pomocí rotace mikrofonního pole, které bylo zasazeno do rotátoru, a které jsem vždy po přehrání jedné nahrávky posunul. Celkem jsem nastavil 8 poloh natočení pole. Pro signál zabouchnutí dveří jsem volil úhly  $0^\circ, 8^\circ, 16^\circ, 24^\circ, 34^\circ, 44^\circ, 56^\circ, 70^\circ$  a  $90^\circ$ . Pro signál tříštění skla jsem volil jiné úhly  $0^\circ, 10^\circ, 20^\circ, 30^\circ, 40^\circ, 50^\circ, 60^\circ, 70^\circ$  a  $90^\circ$ , pro zvuk výstřelu z pistole pak  $0^\circ, 6^\circ, 15^\circ, 22^\circ, 37^\circ, 47^\circ, 58^\circ, 70^\circ$  a  $85^\circ$ . Různé úhly jsem volil kvůli detailnějšímu zpracování, na druhou stranu tato volba lehce znepřehledňuje grafy a výsledky. [1]

Tab. 2.3: Přehled jednotlivých mikrofونů v poli a jejich přesnost. Metoda GCC, kruhové pole, signál bouchnutí dveří, měřený úhel byl  $24^\circ$ . Červený podklad u čísla značí nejvyšší hodnotu naměřeného úhlu, oranžový potom nejmenší. Průměr v modrém poli byl počítán z absolutní hodnoty úhlů.

NaN	22,01	21,27	15,69	11,74	-26,98	-22,03	-23,40	Pohyb ve °
-22,01	NaN	20,98	0,00	-17,33	-23,45	-22,62	-23,52	24,00
-21,27	-20,98	NaN	-12,05	-22,03	-23,19	-22,83	-23,30	
-15,69	0,00	12,05	NaN	-24,66	-20,78	-23,30	-23,34	
-11,74	17,33	22,03	24,66	NaN	-22,01	-22,34	-21,13	
19,20	23,45	23,19	20,78	22,01	NaN	-18,41	-6,26	33,08
22,03	22,62	22,83	23,30	22,34	33,08	NaN	12,05	-26,98
23,40	23,52	23,30	23,34	21,13	15,91	-12,05	NaN	20,63

Příklad zjišťování úhlů směru přicházejícího zvuku můžeme vidět v tab. 2.3. Jde vidět, že obyčejným průměrováním výsledných úhlů z mikrofونu je možno dostat velice nepřesný výsledek, a to kvůli mikrofونům, které nevyhodnotili signál správně. Zlepšení výsledků by mohlo přinést váhování výsledků, popřípadě počítání průměru z vybraných hodnot (např. počítání dalšího průměru z hodnot kolem průměru do padesáti procent). V tab. 2.7 nalezneme vyhodnocení jednotlivých polí a metod. Jedná se o průměry z absolutních hodnot úhlů všech mikrofونů v poli, protože funkce, ze které počítáme úhel, při prohození vstupních signálů vypočítá stejnou hodnotu, ale s jiným znaménkem. [1]

Oproti bakalářské práci je tato práce zaměřena na zjišťování úhlu z malých počtů

vzorků. Postupně jsou zde popsány výpočty na jednotlivých segmentech signálu. Segmenty jsou o délkách 50, 5 a 3 vzorky, přičemž celková délka signálu byla také měněna. Další změnou oproti bakalářské práci je filtrování signálu (omezení signálu na nízkých a vysokých kmitočtech horní a dolní propustí). Všechny skripty byly uzpůsobeny pro různý počet mikrofónů.

Výsledky mikrofonních polí v laboratoři						
Bouchnutí dveří						
Pohyb ve °	Kruhové pole			Křížové pole		
	Metoda GCC	Metoda PHAT	Metoda ML	Metoda GCC	Metoda PHAT	Metoda ML
8	7,47	7,78	7,78	8,46	8,54	8,54
16	13,94	13,98	13,98	14,69	15,20	15,20
24	20,63	20,48	20,48	20,75	20,99	20,99
34	27,12	27,56	27,56	29,43	27,83	27,83
44	35,15	34,96	34,96	34,84	34,91	34,91
56	44,83	47,64	47,64	45,01	43,38	43,38
70	51,46	52,22	52,22	54,29	54,64	54,64
90	61,41	62,13	62,13	60,66	61,74	61,74

Výsledky mikrofonních polí v laboratoři						
Rozbití skla						
Pohyb ve °	Kruhové pole			Křížové pole		
	Metoda GCC	Metoda PHAT	Metoda ML	Metoda GCC	Metoda PHAT	Metoda ML
10	9,42	9,20	9,20	9,46	9,52	9,52
20	16,68	16,97	16,97	17,53	17,25	17,25
30	24,57	24,44	24,44	25,54	25,86	25,86
40	31,98	32,28	32,28	33,37	33,64	33,64
50	39,68	39,34	39,34	39,61	40,21	40,21
60	48,66	49,89	49,89	49,33	49,45	49,45
70	51,62	52,06	52,06	54,66	54,63	54,63
90	61,30	61,63	61,63	61,67	61,97	61,97

Výsledky mikrofonních polí v laboratoři						
Výstřel						
Pohyb ve °	Kruhové pole			Křížové pole		
	Metoda GCC	Metoda PHAT	Metoda ML	Metoda GCC	Metoda PHAT	Metoda ML
6	5,55	5,99	5,99	6,54	6,63	6,63
15	13,05	13,67	13,67	13,63	13,71	13,71
22	18,38	19,23	19,23	19,85	19,24	19,24
37	29,92	29,47	29,47	30,88	30,95	30,95
47	36,87	36,43	36,43	37,72	38,39	38,39
58	48,72	48,86	48,86	45,09	45,50	45,50
75	49,65	49,60	49,60	51,04	53,83	53,83
85	59,85	59,93	59,93	59,19	59,61	59,61

Obr. 2.7: Tabulky výsledků měření v laboratoři, přehled jednotlivých metod [1]



## Metoda GCC

Metoda GCC měla při měření se zvukem zabouchnutí dveří (převažují nízké kmitočty) nejlepší výsledky s křížovým polem při odpovídajícím úhlu  $8^\circ$ , při větším úhlu ( $24^\circ$ ) naopak vychází lépe kruhové pole. Je zajímavé, že při počítání směru pohybu na méně vzorcích je výpočet přesnější. Což je nejspíš dáno tím, že algoritmus spočítá (při přechodu na z 50 vzorků na 5) desetkrát více kombinací možností. Při signálu, ve kterém převažují nižší kmitočty, metoda GCC vyšla nejlépe ze PHAT i ML. Nejmenší počet celkových vzorků, na kterém se počítá směr přicházejícího úhlu u podobného signálu, by neměl být menší než 1000 vzorků, potom narůstá významným způsobem chyba (větší než 16 %).

Pokud má signál široké spektrum (jako například rozbití skla), metoda GCC vychází nejlépe v kruhovém poli. Na tomto druhu signálu je velmi přesná, a to i při malém počtu vzorků zkoumaných vzorků (5 a 3 vzorky). Celková délka signálu při tomto typu signálu by neměla být menší než 100 vzorků, což oproti jiným vstupním signálům je značný rozdíl.

Kruhové pole, zabouchnutí dveří					Křížové pole, zabouchnutí dveří				
50 vzorků/2000					50 vzorků/2000				
	GCC	PHAT	ML			GCC	PHAT	ML	
8,00	7,04	4,21	6,22		8,00	7,80	7,27	6,68	
50 vzorků/20000					50 vzorků/20000				
	GCC	PHAT	ML			GCC	PHAT	ML	
24,00	20,81	20,67	20,51		24,00	20,32	19,06	18,78	
5 vzorků/2000					5 vzorků/2000				
	GCC	PHAT	ML			GCC	PHAT	ML	
8,00	7,17	6,21	7,66		8,00	7,93	7,02	7,14	
3 vzorků/1000					3 vzorků/1000				
	GCC	PHAT	ML			GCC	PHAT	ML	
24,00	19,82	20,67	20,51		24,00	20,26	19,06	18,78	
3 vzorků/100					3 vzorků/100				
	GCC	PHAT	ML			GCC	PHAT	ML	
24,00	18,91	6,79	7,34		24,00	18,27	12,74	13,67	
3 vzorků/42					3 vzorků/42				
	GCC	PHAT	ML			GCC	PHAT	ML	
24,00	18,91	6,79	7,34		24,00	8,29	3,24	3,40	

Obr. 2.8: Výsledky naměřeného úhlu metodami GCC, PHAT a ML v kruhovém a křížovém poli. Typ signálu: zabouchnutí dveří.

Posledním testovaným signálem byl výstřel z pistole. Hlavním problémem takového typu signálu je, že má krátkou dobu trvání. Metoda GCC byla nejlepší při rozestavení mikrofونů v křížovém poli, nicméně při detekování pohybu ve větším úhlu byla chyba znatelně větší. Opět by celková délka signálu neměla být kratší než 1000 vzorků.

## Metoda PHAT

Metoda PHAT se nejvíce jeví jako vhodná metoda ke zkoumání DOA signálu na nízkých kmitočtech. Při pohybu o  $8^\circ$  lépe vychází rozmístění mikrofونů do křížového pole, při pohybu o více stupňů naopak kruhové. Při zkoumání pohybu na malém počtu vzorcích se nejvíce projevilo, jaká byla délka signálu. Minimální délka signálu při zkoumání signálu se opět jeví jako 1000 vzorků.

Výpočet směru příchodu signálu rozbití skla byl mnohem přesnější u malého posunu, kde lépe vyšlo kruhové pole, nicméně rozdíly mezi poli jsou minimální rozdíly, s tím, že se výsledek moc nelišil při počítání na malém počtu vzorků. Při větším úhlu je vhodnější křížové pole. Nejmenší délka signálu s přijatelnou chybou je 100 vzorků.

U zvuku s krátkým trváním je metoda PHAT srovnatelná s metodou GCC.

Kruhové pole, rozbití skla					Křížové pole, rozbití skla				
50 vzorků/2000					50 vzorků/2000				
Stupně	GCC	PHAT	ML		Stupně	GCC	PHAT	ML	
10,00	9,88	9,38	9,71		10,00	8,88	9,65	9,37	
50 vzorků/20000					50 vzorků/20000				
Stupně	GCC	PHAT	ML		Stupně	GCC	PHAT	ML	
30,00	25,77	20,67	24,12		30,00	24,38	23,71	22,70	
5 vzorků/2000					5 vzorků/2000				
Stupně	GCC	PHAT	ML		Stupně	GCC	PHAT	ML	
10,00	9,88	9,29	9,71		10,00	10,22	9,38	8,35	
3 vzorky/1000					3 vzorky/1000				
Stupně	GCC	PHAT	ML		Stupně	GCC	PHAT	ML	
30,00	25,64	20,67	24,12		30,00	24,12	23,71	22,70	
3 vzorky/100					3 vzorky/100				
Stupně	GCC	PHAT	ML		Stupně	GCC	PHAT	ML	
30,00	24,27	21,65	21,73		30,00	24,12	22,83	22,24	
3 vzorky/42					3 vzorky/42				
Stupně	GCC	PHAT	ML		Stupně	GCC	PHAT	ML	
30,00	19,05	11,38	13,04		30,00	13,18	8,13	8,83	

Obr. 2.9: Výsledky naměřeného úhlu metodami GCC, PHAT a ML v kruhovém a křížovém poli. Typ signálu: rozbití skla.

## Metoda ML

Poslední metodou v této skupině je ML. Úhly naměřené vůči prvnímu a druhému mikrofonu byly značně nepřesné (u všech typů signálu), proto je bylo nutné vyřadit z výsledků (zhruba 8 kombinací). Výběr jsem zvolil takový, že jsem z průměru všech naměřených úhlů udělal střední hodnotu  $x$  intervalu  $\langle x - 30^\circ, x + 30^\circ \rangle$  a následně jsem počítal druhý průměr z výsledků, do které byly v tomto intervalu. Metoda ML vykazovala nejlepší výsledky při uskupení mikrofonů v kruhovém poli (při zkoumaném signálu zabouchnutí dveří). Stále platí, že pokud máme dostatečně dlouhý signál, minimálně 1000 vzorků, tak zkoumání DOA na 3 vzorcích má lepší výsledky.

U signálu rozbití skla měla metoda ML spíše průměrné výsledky, nejvhodnější bylo kruhové pole, přičemž bylo nutné mít signál alespoň o délce 1000 vzorků.

Výstřel z pistole zaznamenala DOA metoda nejlépe v kruhovém poli a vyšla nejlépe ze všech.

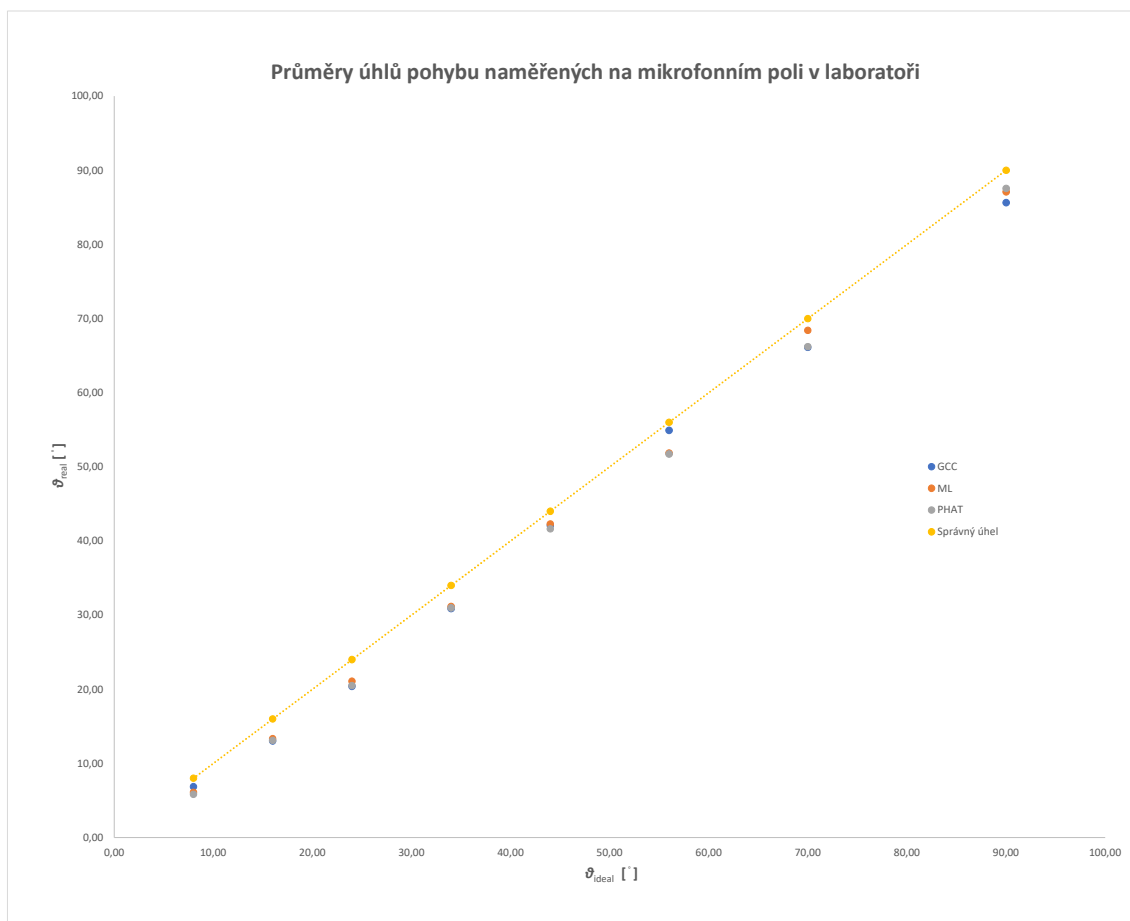
Kruhové pole, výstřel.				
50 vzorků/20000				
Stupně	GCC	PHAT	ML	
6,00	3,98	3,48	5,09	
50 vzorků/20000				
Stupně	GCC	PHAT	ML	
22,00	18,85	19,00	19,15	
5 vzorků/2000				
Stupně	GCC	PHAT	ML	
6,00	5,95	5,78	6,74	
3 vzorky/1000				
Stupně	GCC	PHAT	ML	
22,00	18,85	19,00	19,15	
3 vzorky/100				
Stupně	GCC	PHAT	ML	
22,00	16,57	17,04	19,14	
3 vzorky/42				
Stupně	GCC	PHAT	ML	
22,00	6,48	4,72	5,84	

Křížové pole, výstřel				
50 vzorků/20000				
Stupně	GCC	PHAT	ML	
6,00	6,91	6,45	6,26	
50 vzorků/20000				
Stupně	GCC	PHAT	ML	
22,00	20,27	19,10	18,68	
5 vzorků/2000				
Stupně	GCC	PHAT	ML	
6,00	7,02	5,89	5,83	
3 vzorky/1000				
Stupně	GCC	PHAT	ML	
22,00	20,04	19,10	18,68	
3 vzorky/100				
Stupně	GCC	PHAT	ML	
22,00	20,04	18,83	18,63	
3 vzorky/42				
Stupně	GCC	PHAT	ML	
22,00	10,74	4,31	5,27	

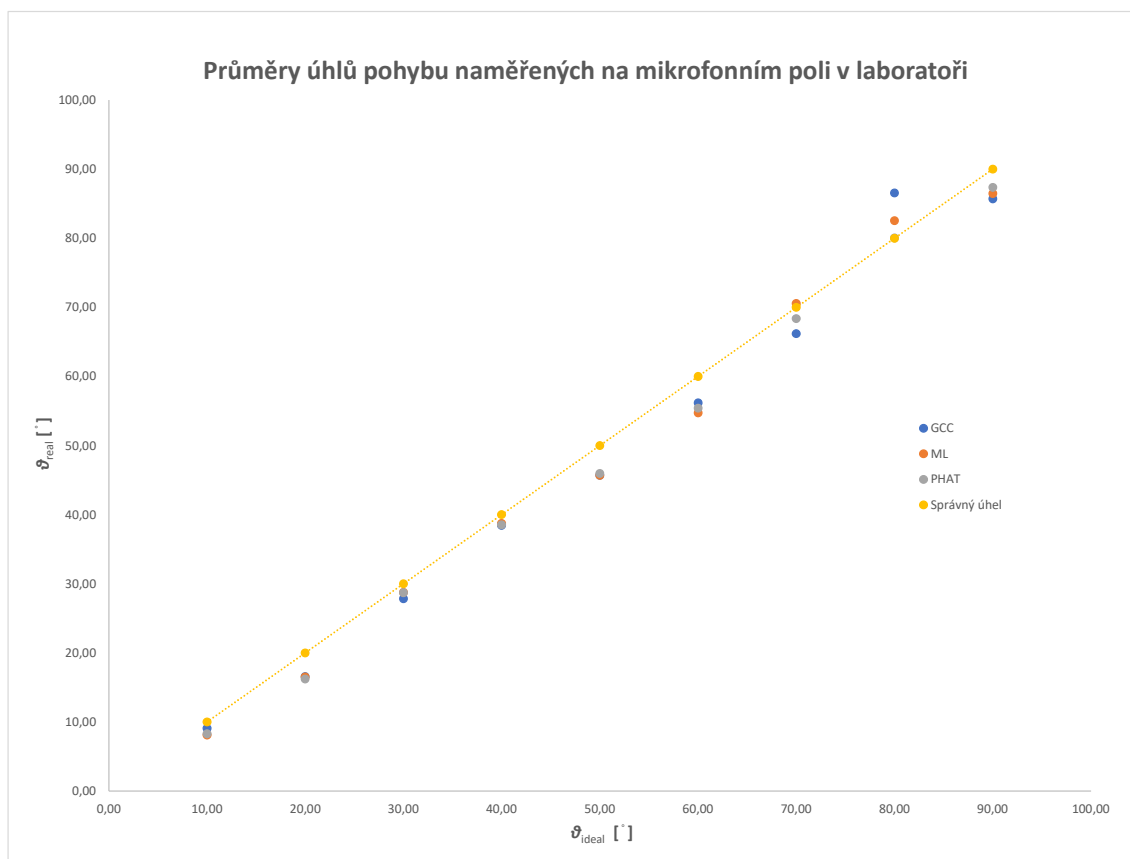
Obr. 2.10: Výsledky naměřeného úhlu metodami GCC, PHAT a ML v kruhovém a křížovém poli. Typ signálu: výstřel.

## Křížové pole

Jak bylo zmíněno výše, výsledky křížového pole jsou velice přesné. Křížové pole poskytuje mnoho různých natočení dvojic mikrofonů, což je oproti kruhovému poli výhoda a má také mnoho kombinací vzdáleností mikrofonů od sebe. Při testování tohoto pole jsem upravil výběr úhlů z výsledné matice. Z naměřených úhlů jsem vzal medián a následně jsem jej považoval jako střední hodnotu intervalu, ve kterém se musely nacházet výsledné úhly, jinak nebyly započítány do celkového průměru. Je to obdobný výběr jako v 2.2.2, s tím, že omezuje chybu výsledků, které jsou o mnoho vyšší než zbylé výsledky. Tento výběr vychází z předpokladu, že extrémní výkyvy (jak malé, tak i velké hodnoty) budou mít podobnou četnost. Ve výsledných tabulkách je brán signál o velikosti 1024 vzorků, přičemž velikost části signálu, ze které se počítal úhel, byl 4 vzorky.

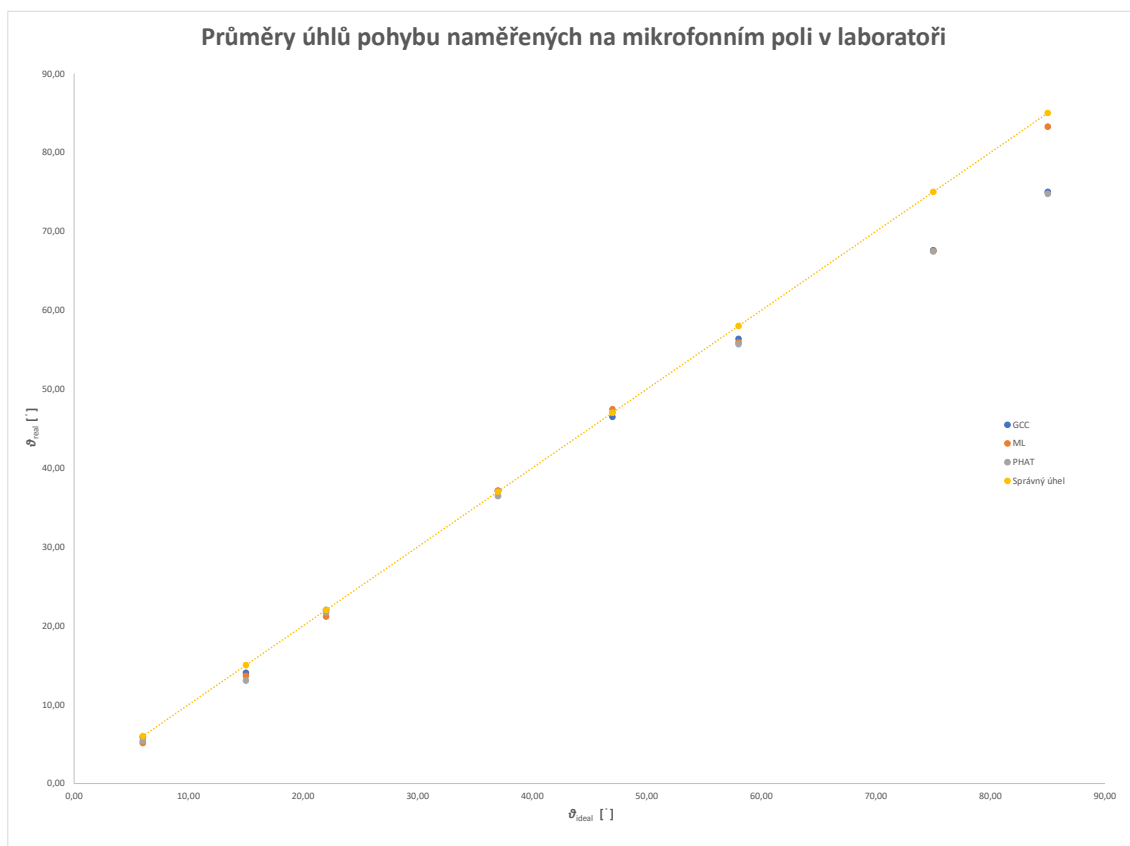


Obr. 2.11: Výsledky průměrů naměřeného úhlu v křížovém poli. Typ signálu: dveře. Velikost signálu 1024 vzorků, úhel počítán na 4 vzorkách



Obr. 2.12: Výsledky průměrů naměřeného úhlu v křížovém poli. Typ signálu: sklo. Velikost signálu 1024 vzorků, úhel počítán na 4 vzorkách.

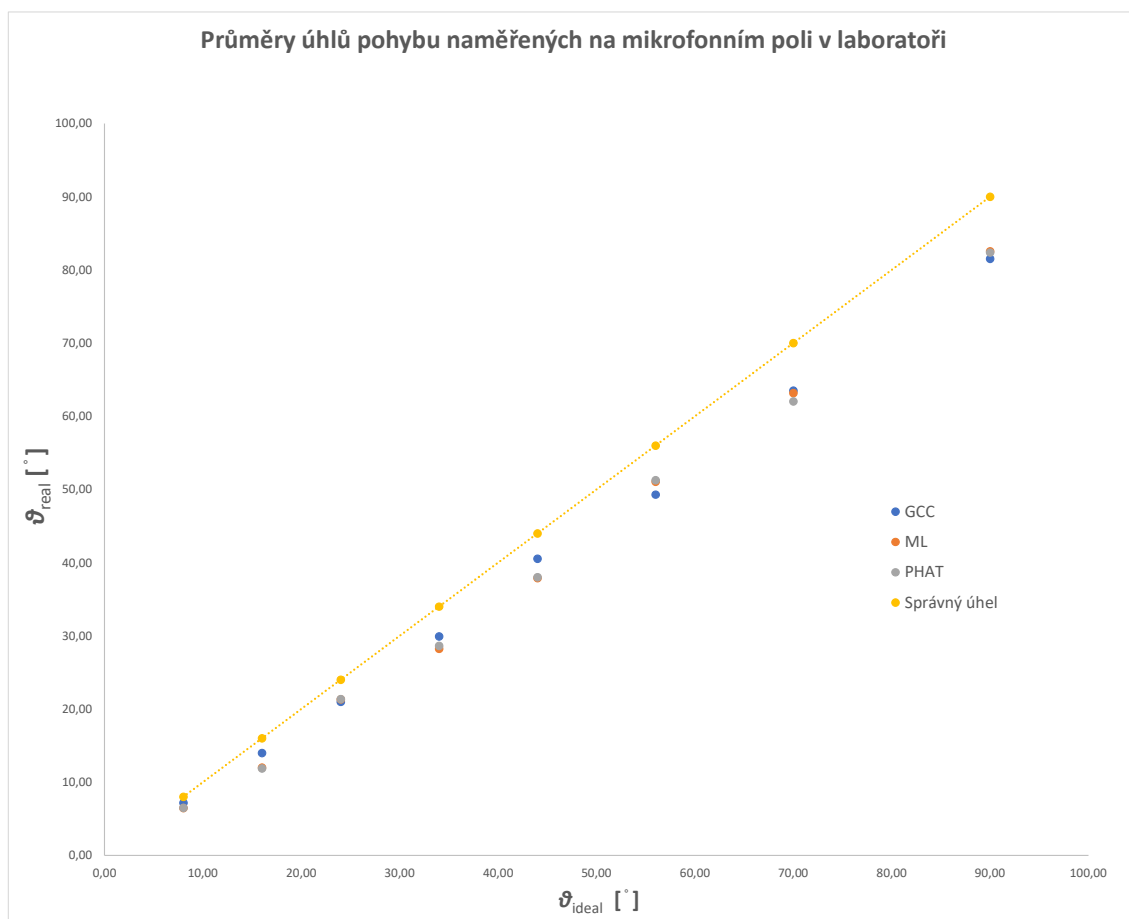
Nutností takového výběru je, zvolit správný interval hodnot, které se budou započítávat do výsledku. Můj výběr byl takový, že jsem bral z menších úhlů vyšší procentuální část např. když medián výsledků vyšel  $8,2^\circ$ , vzal jsem interval 50 % z hodnoty mediánu (v tomto případě by takový interval byl  $\langle 4,1^\circ, 12,3^\circ \rangle$  zatímco při mediánu  $40,8^\circ$  jsem vzal 25 %, tj. interval  $\langle 30,6^\circ, 51^\circ \rangle$ . Průměrná chybovost jednotlivých mikrofonních párů byla v křížovém poli následující: u GCC bylo vynecháno průměrně 20 párů, u PHAT bylo vynecháno 22 párů, chybovost ML byla 27 párů. Přičemž tato chybovost se zvedala zároveň s narůstajícím správným úhlem (tato charakteristika byla potvrzena již v bakalářské práci a platí i pro malé počty vzorků).



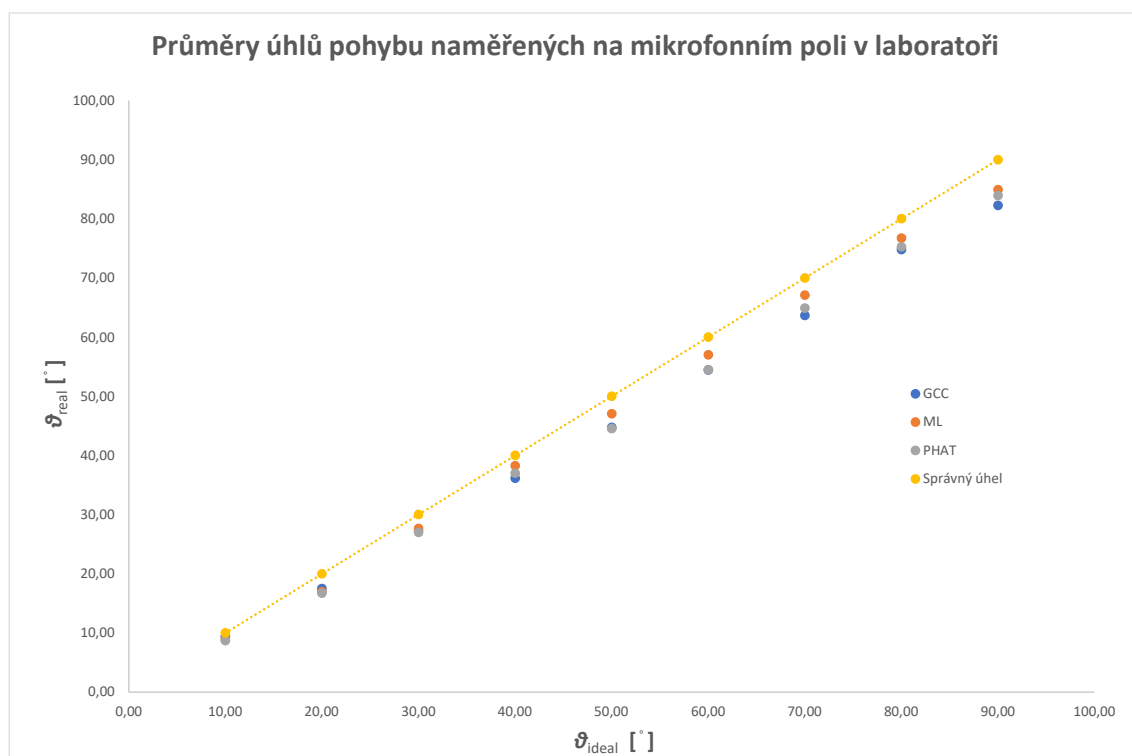
Obr. 2.13: Výsledky průměrů naměřeného úhlu v křížovém poli. Typ signálu: výstřel. Velikost signálu 1024 vzorků, úhel počítán na 4 vzorkách.

## Kruhové pole

Výběr úhlů z výsledné matice probíhal stejně jako u kruhového pole. Velkým rozdílem je, že se velice zmenšila chybovost párů mikrofonů (větší vzdálenosti mezi mikrofony jsou přesnější). Většinou tyto chyby byly na stejných párech mikrofonů, dalo by se uvažovat i o celkovém vynechání takových párů. Je také možné si všimnout, že metoda ML má při větších úhlech větší přesnost. Je to dáno nejspíše tím, že již od začátku část párů mikrofonů (hlavně v kombinaci s prvním mikrofonem) od začátku vykazovala hodnoty kolem  $100^\circ$ , které se do výsledků při malém úhlu nezapočítaly. Avšak při úhlu  $90^\circ$  je již takový výsledek započten. Opět by bylo objektivnější takové páry mikrofonů nezahrnovat do výsledku, nicméně při pokusu mimo laboratoř tyto páry fungují. Z tohoto důvodu jsem je do výsledků zahrnul.

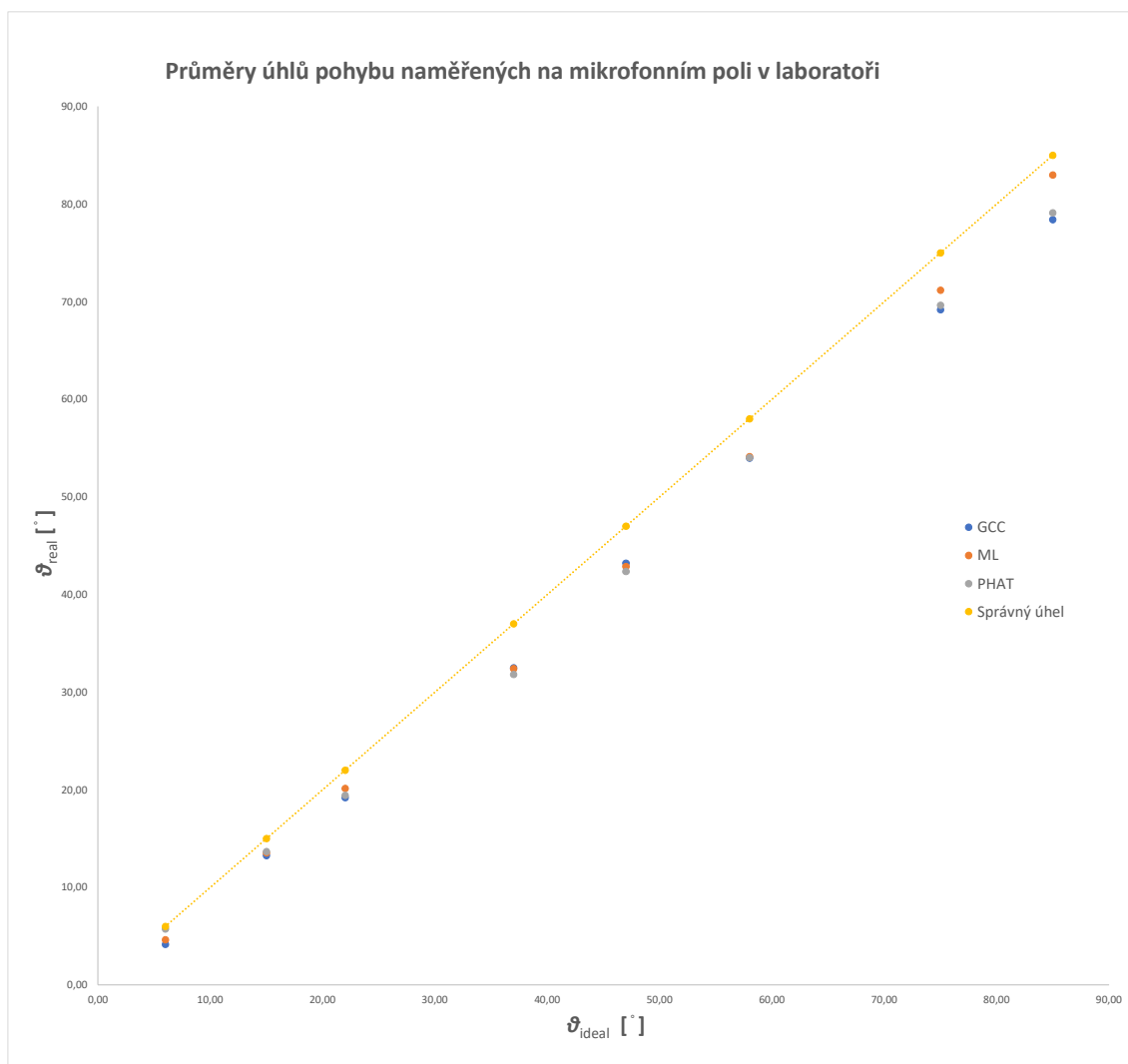


Obr. 2.14: Výsledky průměrů naměřeného úhlu v kruhovém poli. Typ signálu: dveře. Velikost signálu 1024 vzorků, úhel počítán na 4 vzorkách



Obr. 2.15: Výsledky průměrů naměřeného úhlu v kruhovém poli. Typ signálu: rozbití skla. Velikost signálu 1024 vzorků, úhel počítán na 4 vzorkách

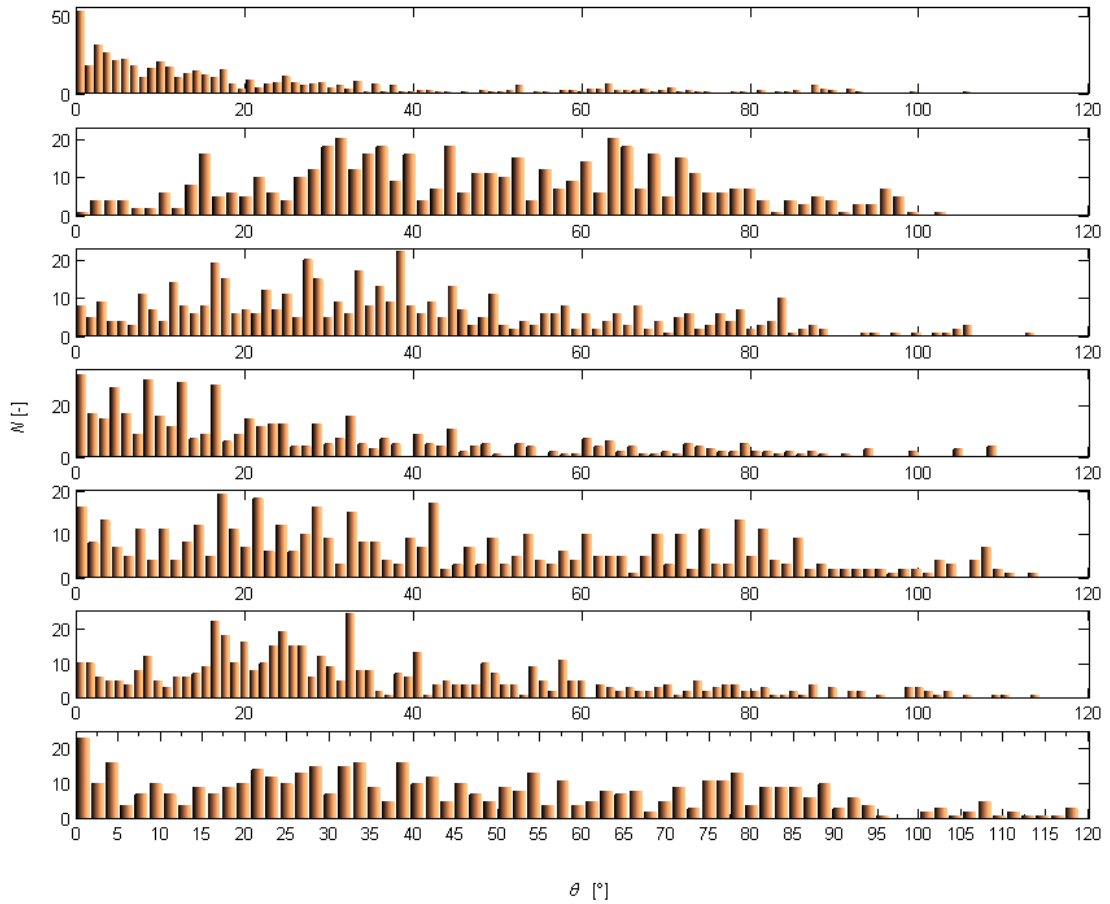




Obr. 2.16: Výsledky průměrů naměřeného úhlu v kruhovém poli. Typ signálu: výstřel. Velikost signálu 1024 vzorků, úhel počítán na 4 vzorkách

## Energetická metoda

Energetická metoda bohužel nedává jednoznačné výsledky, a to ani při přepočítávání ze formátu signálu čtyř mikrofónů 0. řádu, ani při přepočtu z A-formátu signálu. Nejpravděpodobnější „správné“ výsledky dával A-formát signálu, kdy vstupním signálem byl růžový šum omezený na kmitočtový rozsah mikrofónu, úhel byl měřený nivelačním laserem (chyba určení úhlu by měla být do  $2^\circ$ ). Při průměrování směrů úhlů vyšel správně pouze první úhel, a to při posunu o  $10^\circ$ . Distribuční funkci výsledné matice je na obr. 2.17.



Obr. 2.17: Četnosti nalezených úhlů výsledné matice energetické metody, typ signálu: růžový šum, A-formát signálu. Shora: pohyb o  $10^\circ$ ,  $20^\circ$ ,  $30^\circ$ ,  $40^\circ$ ,  $50^\circ$ ,  $60^\circ$ ,  $70^\circ$

### 2.2.3 Výsledky z měření v reálném prostředí

Měření v reálném prostředí proběhlo pouze s šesti mikrofony, a to kvůli nedostupnosti zvukové karty s více vstupy. Mikrofony jsem rozestavěl do křížového pole, jelikož dávalo jedny z nejlepších výsledků, s tím, že mikrofony číslo 6 a 7 2.3 jsem vynechal. Testovací signály byly: zvuk ze sirény, generátoru zvuku rozbití skla a startovací pistole. Měření probíhalo v prostředí, kde ve vzdálenosti cca. 8 metrů byla cihlová zeď, tudíž v záznamu byl silný zvuk z odrazu. Dalšími ručovými aspekty byly: vítr, zpěv ptáků, lidé atd. Jednalo se tedy o velmi rušné prostředí. Teplota vzduchu byla okolo 30 °C. Výsledky z nahrávání jsou zaznamenány v tab. 2.4. [1]

Po dohodě s vedoucím práce byla provedena čtyři měření (tři výsledky) u výstřelu z nedostatku nábojů. Při zaznamenávání zvuku sirény celý pokus narušil nepřiměřený hluk, tudíž jsem byl nucen přerušit měření a byla použita pouze čtyři měření. Měření s generátorem zvuku rozbití skla proběhlo bez problémů. Celkově však nahrávání i algoritmy splnily účel a fungují. [1]

Tab. 2.4: Výsledky měření v reálném prostředí, přehled jednotlivých metod. [1]

Výsledky mikrofonních polí v reálném prostředí						
Generátor rozbití skla				Siréna		
Křížové pole				Křížové pole		
Pohyb ve °	Metoda GCC	Metoda PHAT	Metoda ML	Metoda GCC	Metoda PHAT	Metoda ML
10	13,36	12,62	12,62	44,86	10,94	10,94
20	26,62	26,46	26,46	28,55	17,66	17,66
30	36,41	35,67	35,67	29,61	25,25	25,25
40	39,07	44,29	44,29	-	-	-
50	49,84	46,01	46,01	-	-	-
60	50,97	52,25	52,25	-	-	-
70	59,50	60,42	60,42	-	-	-
80	70,54	72,42	72,42	-	-	-

Výstřel			
Pohyb ve °	Metoda GCC	Metoda PHAT	Metoda ML
20	24,84	24,69	24,69
40	0,00	0,00	0,00
60	55,52	60,58	60,58

V diplomové práci jsem se zaměřil na počítání úhlu na málo vzorkách. Oproti výsledkům na více vzorkách nejsou metody tolik přesné, nicméně stále dostačující. Ve výsledných tabulkách je brán signál o velikosti 1024 vzorků, přičemž velikost části signálu, ze které se počítal úhel, byl 4 vzorky.

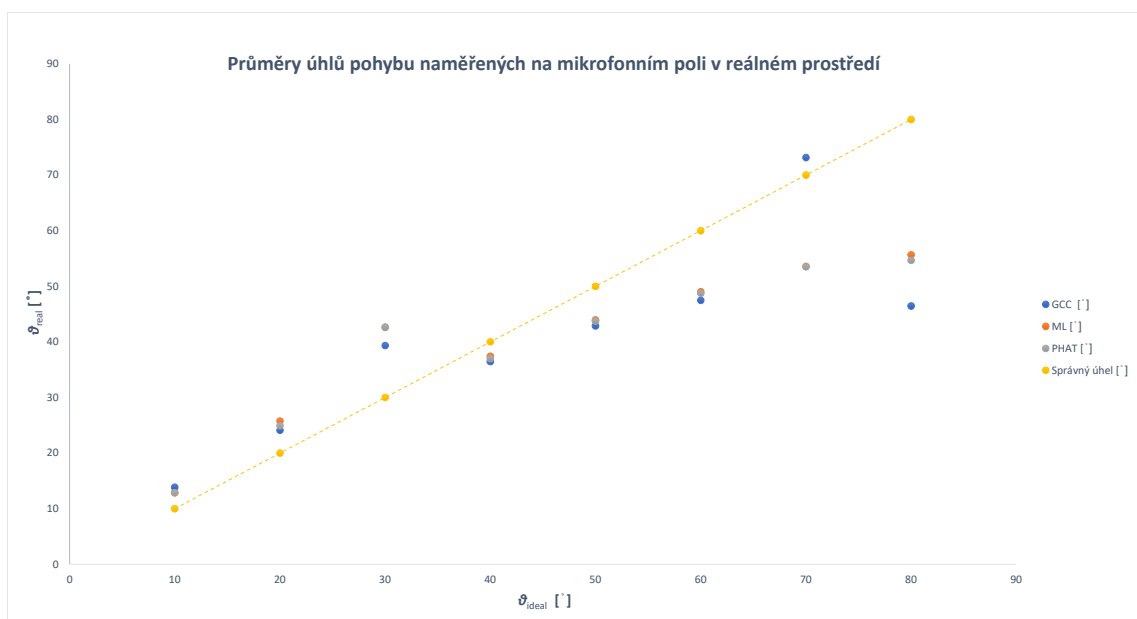
$\vartheta$ [deg.]	GCC	ML	PHAT
10	38,75	3,48	3,07
20	29,6	4,03	4,03
30	25,61	3,99	4,12

Obr. 2.18: Výsledky průměrů naměřeného úhlu v křížovém poli a reálném prostředí. Typ signálu: siréna. Velikost signálu 1024 vzorků, úhel počítán na 4 vzorkách.

$\vartheta$ [deg.]	GCC	ML	PHAT
10	49,63	34,39	34,61
20	0,00	0,00	0,00
30	83,07	29,93	28,65

Obr. 2.19: Výsledky průměrů naměřeného úhlu v křížovém poli a reálném prostředí. Typ signálu: výstřel. Velikost signálu 1024 vzorků, úhel počítán na 4 vzorkách.

Výsledky z reálného prostředí při malém počtu vzorků byly pouze průměrovány, což značně znepřesnilo výsledky. Celkově je interpretace naměřených hodnot problematická a výsledky u typu signálu siréna tab.2.18 a výstřel tab. 2.19 nejsou směrodatné. U výsledků typu signálu rozbití skla obr. 2.20 je možné pozorovat opět větší chybu na vyšších úhlech, což je dáno tím, že celková délka signálu je krátká. Další možností je, zvolit jiný výběr z výsledné matice.



Obr. 2.20: Výsledky průměrů naměřeného úhlu v křížovém poli a reálném prostředí. Typ signálu: rozbití skla. Velikost signálu 1024 vzorků, úhel počítán na 4 vzorkách.

### 3 ZÁVĚR

V diplomové práci jsem se soustředil na souhrn různých metod, kterými se dá odhadnout směr přicházejícího pole. Zahrnul jsem zde i metody, které se nezdají být vhodné pro lokalizaci malým mikrofonním polem, a to kvůli přehlednosti a plnosti tématu. Popsal jsem zde základy metod MEM, křížová korelace, GCC, PHAT, ML, AAR, CPA, MUSIC, energetická metoda, Beamforming (DAS, FAS), ESPRIT. Samozřejmě existuje mnoho dalších metod, které jsem nepopsal, mnoho jich je ve zdrojích. Jsou zde popsány faktory, které nejvíce ovlivňují výsledky měření, ať už věrnost snímání, filtraci signálu atd. Dále v teoretické části rozebírám matematické postupy, např. STFT, která je jádrem energetické metody, či vzájemnou korelaci, která se uplatní u GCC, PHAT a ML.

Ze začátku praktické části se věnuji návrhu držáku mikrofonního pole, které jsem pro lepší výsledky upravil viz 2.2. Bohužel takto navržené pole nemohlo být zkonstruováno kvůli pandemii viru Covid-19 (byl vyhlášen nouzový stav země). Používal jsem tedy nahrané soubory z bakalářské práce, což mělo své výhody i nevýhody. Výhodou bylo, že díky různým inovacím se mi oproti bakalářské práci povedlo značně zpřesnit výsledky hledání směru příchodu zvuku (laboratorní nahrávky jsou přesné a validní) a je možné výsledky tudíž bez problému porovnávat. Velkou nevýhodou je, že neproběhlo korekční nahrávání zvuků v reálném prostředí, protože interpretace výsledků v reálném prostředí z bakalářské práce je problematická (nedošlo k měření bez omezení).

Dále jsou podrobně popsány podrobně naprogramované funkce v Matlabu. Následují výsledky jednotlivých funkcí. Jsou zde popsány možnosti výběru z výsledné matice úhlů. Nejlépe vyšlo určení mediánu s následným výběrem úhlů z výsledné matice v určitém intervalu se středem v mediánu. Co se týče metod GCC, PHAT a ML, tak výsledky vycházely podobně, avšak nejlépe vyšla ML. Oproti bakalářské práci se mi povedlo odlišit výsledky metody ML a PHAT tím, že jsem celý signál zpracovával předem (horní i dolní propustí, vybíráním části signálu). Naprogramované funkce bez větších chyb fungovaly i na malém počtu vzorků (4 vzorky k určení úhlu stačí), avšak za podmínky, že celková délka signálu nebude kratší než 1000 vzorků. Jediný problém byl u metody ML, kdy určitá část dvojic mikrofónů vykazovala konstantní úhel. Energetická metoda bohužel nedává jednoznačné výsledky. Je složité odhadovat, kde nastala chyba.

Rozdíly mezi poli byly též minimální, lepším polem pro měření vychází křížové pole. Výhodou tohoto pole je, že existuje více kombinací vzdáleností, tudíž se zlepšuje při správném výběru přesnost. Jistou nevýhodou je, že pokud se zvolí špatná metoda výběru či špatný interval, chybovost dvojic mikrofónů při určování příchodu směru zvuku vzroste. Kruhové pole oproti křížovému mělo přibližně konstantní chybu ur-

čení směru, avšak určovalo výsledek spíše pod správnou hodnotou.

Použití algoritmu při velmi malých časových diferencích signálu v reálném prostředí se dá odhadnout jako úspěšný, ačkoliv výběr z matice byl vybírán pouhým průměrem z naměřených úhlů. Od úhlu cca.  $50^\circ$  nebyl výsledek přesný, ale vzhledem k tomu, že byl úhel počítán na signálu o délce 1024 vzorků, je nepravděpodobné, že by se zdroj zvuku pohyboval rychleji než právě  $50^\circ$  za 0,023 s (při vzorkovacím kmitočtu 44,1 kHz).

Další práci na této práci by mohlo být např. určování roviny, odkud jde zvuk (což bylo v plánu již v této práci, nicméně nebylo možno nahrát referenční signál, a to kvůli výše zmiňované pandemii), dále zjišťování elevace příchodu zvuku, či naprogramování dalších algoritmů.

# LITERATURA

- [1] KUBIŠTA, Ladislav. *Mikrofonní pole pro detekci směru přicházejícího zvuku: bakalářská práce*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací, 2016. 53 s. Vedoucí práce byl Ing. Jiří Schimmel, Ph.D.
- [2] SCHIMMEL, Jiří. *Elektroakustika* [online]. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací, Purkyňova 118, 612 00, Brno, 2016 [cit. 7. 12. 2016]. ISBN 978-80-214-4716-5. Dostupné z: <[https://www.vutbr.cz/www\\_base/priloha.php?dpid=71624](https://www.vutbr.cz/www_base/priloha.php?dpid=71624)>.
- [3] KHADDOUR, Hassan. *A Comparison of Algorithms of Sound Source Localization Based on Time Delay Estimation*. Elektorevue [online]. Department of Telecommunications FEEC, Brno University of Technology, Purkyňova 118, 61200, Brno, 2011, 1–7 [cit. 2. 12. 2016]. ISSN 1213–1539. Dostupné z:<<http://elektorevue.cz/en/articles/analogue-technics/90a-comparison-of-algorithms-of-sound-source-localization-based\discretionary{-}{-}{-}on-time-delay-estimation-1-1-1/>>>.
- [4] CARTER, G. C. *Tutorial Overview of Coherence and Time Delay Estimation*, in *Coherence and Time Delay Estimation—An Applied Tutorial for Research, Development, Test, and Evaluation Engineers*, vol. 1, 1993, pp. 1–27
- [5] KNAPP C. and G. CARTER. *The generalized correlation method for estimation of time delay*, in *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 24, no. 4, pp. 320-327, Aug 1976. doi: 10.1109/TASSP.1976.1162830
- [6] VLACH, J., Z. PRŮŠA a P. RAJMIC. *Interaktivní applety pro vybrané operace se signály*. Elektorevue. Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací, Purkynova 118, 612 00 Brno, Česká republika, 2009, , 1–13. ISSN 1213–1539. Dostupné také z:<<http://elektorevue.cz/cz/clanky/zpracovani-signalu/0/interaktivni-applety-pro-vybrane-operace-se-signaly-1/>>>.
- [7] HLEDÍK, Stanislav. *Fourierova transformace* [online].[cit. 04. 12. 2016]. Ústav fyziky Filozoficko-přírodovědecké fakulty Slezské univerzity v Opavě.
- [8] ADRIAENSEN, Fons. *A Tetrahedral Microphone Processor for Ambisonic Recording*. Parma, Italy: Laboratorio di Acustica ed Elettroacustica. Dostupné z:<<https://kokkinizita.linuxaudio.org/papers/tetraproc.pdf>>.



- [9] SANTAMARÍA, Ignacio a Javier VÍA. *Estimation of the magnitude squared coherence spectrum based on reduced-rank canonical coordinates* [online]. University of Cantabria, Santander, 39005, Spain, ©2007 [cit. 04.12.2016]. Dostupné z: <<https://pdfs.semanticscholar.org/379b/096a034a0c4a87a78dc10c9baa4668cd2ff8.pdf>>.
- [10] HECKBERT, Paul. *Computer Graphics 2: Fourier Transforms and the Fast Fourier Transform (FFT) Algorithm* [online]., 1–13 [cit. 5.12.2016]. Dostupné z: <<https://www.cs.cmu.edu/afs/andrew/scs/cs/15-463/2001/pub/www/notes/fourier/fourier.pdf>>.
- [11] KHADDOUR, Hassan a Jiří Schimmel. *Multiple Sound Sources Localization Using Energetic Analysis Method*. Elektrorevue. [online]. Department of Telecommunications FEEC, Brno University of Technology, Purkyňova 118, 61200, Brno, 2012, 1–9 [cit. 25.05.2017]. ISSN 1213–1539. Dostupné z: <<http://elektrorevue.cz/en/articles/analogue-technics/0/multiple-sound-sources-localization-using-energetic-analysis-method/>>.
- [12] NURUZZAMAN, A., O. BOYRAZ a B. JALALI. *Time-Stretched Short-Time Fourier Transform*. *IEEE Transactions on Instrumentation and Measurement* 2006, 55(2), 598-602. DOI: 10.1109/TIM.2006.864246. ISSN 0018-9456. [cit. 25.05.2017] Dostupné také z: <<http://www.photonics.ucla.edu/media/publication/journal/101109TIM2006864246.pdf>>
- [13] HEINZEL, G., A. RÜDIGER a R. SCHILLING. *Spectrum and spectral density estimation by the Discrete Fourier transform (DFT), including a comprehensive list of window functions and some new at-top windows* [online]., 1–84 [cit. 25.05.2017]. Dostupné z: <<http://hdl.handle.net/11858/00-001M-0000-0013-5579-7>>
- [14] *Short-time Fourier transform* In: Wikipedia: the free encyclopedia [online], San Francisco (CA): Wikimedia Foundation, 2001- [cit. 30.05.2017]. Dostupné z: <[https://en.wikipedia.org/wiki/Short-time\\_Fourier\\_transform](https://en.wikipedia.org/wiki/Short-time_Fourier_transform)>
- [15] KURC, David *Identifikace zdrojů hluku pomocí beamformingu*: diplomová práce. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací, 2011. 74 s. Vedoucí práce byl Ing. Jiří Schimmel, Ph.D.
- [16] ZOU, Dong-yao, Bi-wei LIU a Wei YANG. *An Efficient TDOA and GROA Localization Mechanism Based on GSO Algorithm with Flight Theory*. *Journal of Computers* [online]. 2017, 2017, 28(5), 11-25 [cit. 2020-04-09]. DOI:

- 10.3966/199115992017102805002. Dostupné z: <[http://www.csroc.org.tw/journal/J0C28\\_5/1991-1599-28.5-02.pdf](http://www.csroc.org.tw/journal/J0C28_5/1991-1599-28.5-02.pdf)>
- [17] TUNCER, T. Engin a Benjamin FRIEDLANDER, ed. *Classical and modern direction-of-arrival estimation*. Boston: Academic, c2009, 456 s. DOI: <<https://doi.org/10.1016/C2009-0-19135-3>>. ISBN 01-237-4524-1.
  - [18] CAMERON, Karleigh J. *FDOA-based passive source localization: a geometric perspective*. Colorado, 2018. Doctoral dissertation. Colorado State University. Vedoucí práce Dan Bates.
  - [19] Li, D., Hu, Y.H. *Energy-Based Collaborative Source Localization Using Acoustic Microsensor Array*. EURASIP J. Adv. Signal Process. 2003, 985029 (2003). <<https://doi.org/10.1155/S1110865703212075>>
  - [20] BARABELL, A. J., J. CAPON, D. F. DELONG, R. JOHNSON a K. D. SENNE. *Performance Comparison of Superresolution Array Processing Algorithms, Revised*. Massachusetts: Massachusetts institution of technology Lexington Lincoln laboratory, 1998. AD-a347 296. Dostupné také z:<[https://www.researchgate.net/publication/270218533\\_Performance\\_Comparison\\_of\\_Superresolution\\_Array\\_Processing\\_Algorithms](https://www.researchgate.net/publication/270218533_Performance_Comparison_of_Superresolution_Array_Processing_Algorithms)>
  - [21] SHENG, Xiaohong a Yu-Hen HU. *Maximum likelihood multiple-source localization using acoustic energy measurements with wireless sensor networks*, in IEEE Transactions on Signal Processing, vol. 53, no. 1, pp. 44-53, Jan. 2005. doi: 10.1109/TSP.2004.838930 Dostupné také z:<<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1369649&isnumber=29977>>
  - [22] YANG, Qing, Alvin LIM, Kenan CASEY a Raghu-Kisore NEELISETTI. *An Enhanced CPA Algorithm for Real-Time Target Tracking in Wireless Sensor Networks*. International Journal of Distributed Sensor Networks [online]. 2009, 5(5), 619-643 [cit. 2020-03-25]. DOI: 10.1080/15501320802581318. ISSN 1550-1477. Dostupné z:<<http://journals.sagepub.com/doi/10.1080/15501320802581318>>
  - [23] *MUSIC Super-Resolution DOA Estimation*. MathWorks [online]. 2020 [cit. 2020-04-17]. Dostupné z:<<https://www.mathworks.com/help/phased/ug/music-super-resolution-doa-estimation.html>>
  - [24] HIDRI Adel, Meddeb SOUAD, Abdulqadir ALAQEELI a AMIRI Hamid. *Beamforming Techniques for Multichannel audio Signal Separation*. International Journal of Digital Content Technology and its Applications [online]. 2012, 6(20),

- 659-667 [cit. 2020-04-17]. DOI: 10.4156/jdcta.vol6.issue20.72. ISSN 1975-9339. <Dostupnéz:[https://www.researchgate.net/publication/233982178\\_Beamforming\\_Techniques\\_for\\_Multichannel\\_audio\\_Signal\\_Separation](https://www.researchgate.net/publication/233982178_Beamforming_Techniques_for_Multichannel_audio_Signal_Separation)>
- [25] NOVOA José, Rodrigo MAHU, DÍAZ Alejandro, Jorge WUTH, STERN Richard, Nestor YOMA. (2019). *Weighted delay-and-sum beamforming guided by visual tracking for human-robot interaction*. Speech Processing and Transmion Laboratory [online]. 2019 <Dostupnéz:[https://www.researchgate.net/publication/333866041\\_Weighted\\_delay-and-sum\\_beamforming\\_guided\\_by\\_visual\\_tracking\\_for\\_human-robot\\_interaction](https://www.researchgate.net/publication/333866041_Weighted_delay-and-sum_beamforming_guided_by_visual_tracking_for_human-robot_interaction)>
- [26] ALSHEHHI, Abdulla, Mohammad Luai HAMMADIH, M. Sami ZITOUNI, Saif ALKINDI, Nazar ALI a Luis WERUAGA. *Linear and Circular Microphone Array for Remote Surveillance: Simulated Performance Analysis* [online]. Khalifa University of Science, Department of Electrical and Computer Engineering, 2017, 1-4 [cit. 2020-04-18]. Dostupné z: <[https://www.researchgate.net/publication/291650227\\_Linear\\_and\\_Circular\\_Microphone\\_Array\\_for\\_Remote\\_Surveillance\\_Simulated\\_Performance\\_Analysis](https://www.researchgate.net/publication/291650227_Linear_and_Circular_Microphone_Array_for_Remote_Surveillance_Simulated_Performance_Analysis)>
- [27] YERRAMSETTY, Hemanth. *Microphone Array Wiener Beamformer and Speaker Localization With emphasis on WOLA Filter Bank*. Department of Signal Processing, 2012. Master's Thesis. Blekinge Institute of Technology. Vedoucí práce Dr. Nedelko Grbic. Dostupné z: <<https://www.diva-portal.org/smash/get/diva2:829457/FULLTEXT01.pdf>>

# SEZNAM PŘÍLOH

A Zdrojové výpisy z Matlabu	72
B Obsah přiloženého CD	77

## A ZDROJOVÉ VÝPISY Z MATLABU

Výpis A.1: GCC metoda naprogramovaná v Matlabu.

```
1 function posun0 = faze(x,x1,d,fs)
2     %funkce se vstupními par.
3     %x je 1. signal
4     %x1 je 2. signal
5     %d je vzdalenost v metrech
6     %fs je vzork. kmitocet
7 c = 340; %rychlost zvuku
8 konst = c/d;
9 if(length(x1)>length(x)) %podmínka pro délku matice
10     x1 = x1(1:length(x));
11 else
12     x = x(1:length(x1));
13 end;
14 X1 = fft (x);
15 X2 = fft (x1);
16 Sx1 = (X1).*conj(X2);
17 fx = ifft (Sx1); %výsledná korelace v počtu vzorků
18
19 index_maxima = find(fx == max(fx),1) - 1; %hledání max.
20 if index_maxima > length(fx)/2
21     index_maxima = (index_maxima-length(fx));
22 end;
23 posun = index_maxima/fs;
24 theta = acosd(posun*konst);
25 plot(fx);
26 disp(['index_maxima: ' num2str(index_maxima)]);
27 disp(['posun[s]: ' num2str(posun)]);
28 disp(['uhel[ $^{\circ}$ ]: ' num2str(theta)]);
```

Výpis A.2: Část kódu naprogramované metody PHAT v Matlabu.

```
1 %predchozi kod je stejny jako u GCC
2 X1 = fft (x);
3 X2 = fft (x1);
4 Sx1 = (X1).*conj(X2);
5 Sx2 = abs(Sx1);
6 Fx = Sx1./Sx2;
7 fx = ifft (Fx);
8 %zpetna transformace, nasleduje hledani maxima
```

Výpis A.3: Část kódu naprogramované metody ML v Matlabu.

```
1 %predchozi kod je stejny jako u GCC
2 X1 = fft (x);
3 X2 = fft (x1);
4 Gx12 = (X1).*conj(X2); %ruzna vykonova spektra
5 Gx11 = (X1).*conj(X1);
6 Gx22=(X2).*conj(X2);
7 delta2 = ((Gx12).^2)/(Gx11.*Gx22); %vykonove spektrum
8 Fiml= (1./abs(Gx12)).*(abs(delta2)./(1-abs(delta2)));
9 %vysledny filtr
10 fx = ifft (Fiml.*X1.*conj(X2));
11 %zpetna transformace, nasleduje hledani maxima
```

Výpis A.4: Jednoduché naprogramování umělého zvuku v Matlabu.

```
1 sum = randn (1,110); %matice sumu
2 sum1 = sum;
3 nuly = zeros (1,7); %matice s nulami
4 sum = [sum,nuly]; %kvuli stejne delce matice
5 sum2 = [nuly,sum1]; %posunutí signálu o vzorek nuly
```

### Výpis A.5: Energetická metoda naprogramovaná v Matlabu

```

1 function energetic =
2   Energeticka_metoda (x, posun, nfft, okno, fs)
3   %počítá směr přicházejícího zvuku
4   % x - vstupní signál o čtyřech mikrofonech
5   % okno - délka analýzy Hammingovým oknem
6   % posun - posun o kolik se chci posunout v čase
7   % nfft - počet FFT bodů; fs - vzorkovací frekvence, Hz
8   x1=x(1,:); x2=x(2,:); x3=x(3,:); x4=x(4,:);
9   %sloupcový vektor
10  Xfb=x1+x2; %výpočty jednotlivých osmičkových mikrofónů
11  Ylr = x3+x4;
12  Wc = ((Xfb.^2)+(Ylr.^2)).^(1/2);
13  %výpočet mikrofónu ve středu
14  z0 = 413; %impedance vzduchu
15  Xfb = Xfb(:); %udělá z x sloupcový vektor
16  % délka signálu
17  delkas = length(Xfb);
18  % periodické hammingovo okno
19  blackhar = blackmanharris(okno, 'periodic');
20  % stft určení matice a rozdělení
21  pocetrad = ceil((1+nfft)/2);
22  %spočítá řádky pro stft, ceil zaokrouhluje k +nekonečnu
23  pocetsloupcu = 1+fix((delkas-okno)/posun);
24  % spočítá sloupce pro stft, fix zaokrouhluje dolů
25  Xvysl = zeros(pocetrad, pocetsloupcu);
26  % vytvoří prázdnou matici o délce sloupců a výšce řádků
27  indx = 0;
28  % index času (kolikátý vzorek rozkládám)
29  % STFT
30  for i = 1:pocetsloupcu
31      xw = Xfb(indx+1:indx+okno);
32      %signál x od časového indexu 1 po časový index + okno,
33      windowf = xw.*blackhar;
34      % FFT
35      X = fft(windowf, nfft);
36      Xvysl(:, i) = X(1:pocetrad);
37      %uložení do matice výsledků
38      indx = indx + posun; %posun indexu
39  end

```

Výpis A.6: Energetická metoda naprogramovaná v Matlabu

```

1 [q,w]=size(Wvysl);
2 Ix = zeros(q,w);
3 Iy = zeros(q,w);
4
5 for high = 1:w
6 Ix(:,high)=(1/(power(2,0.5)*z0))*real...
7     ... (Xvysl(:,high).*conj(Wvysl(:,high)));
8
9 Iy(:,high)=(1/(power(2,0.5)*z0))*real...
10    ... (Yvysl(:,high).*conj(Wvysl(:,high)));
11 end
12
13
14 [m,n]=size(Ix);
15 Vyslalfa = zeros(m,n);
16
17 for i = 1:n
18     for j = 1:m
19         if Iy(j,i)>=0
20 Vyslalfa(:,i)=atan((-Iy(:,i)./Ix(:,i)))*180/pi;
21         else
22 Vyslalfa(:,i)=(atan((-Iy(:,i)./Ix(:,i)))-pi)*180/pi;
23         end
24     end
25 end
26
27 krok_fr=1;
28 % počet bodů rozdělení výšky matice (frekvenční složky)
29 [vys,n]=size(Vyslalfa);
30
31 vec = -360:1:360;
32 [~,krok] = size(vec);
33 vys = vys - mod(vys,krok_fr)+1;
34 cetnost=zeros(floor(vys/krok_fr),krok);
35
36 frekp = 1;

```



Výpis A.7: Energetická metoda naprogramovaná v Matlabu

```

1 for time = 1:n
2 for p = 1:(vys-1)
3     Vyslalfap = Vyslalfa(p,time)';
4
5     for k = 1:krok-1;
6         akprvek =floor(Vyslalfap);
7         if (akprvek>=vec(k) && akprvek<=vec(k+1))
8             cetnost(p,k) = cetnost(p,k)+1;
9             break
10        end
11    end
12 end
13 end
14
15 energetic=cell(p,krok);
16
17 for i = 1: krok
18     for k = 1:p
19         cetnostp= cetnost(k,:);
20         [~,energeticp_index] =find(cetnostp==...
21                                     ...max((cetnostp),[ ],2));
22         energetic{k,i} = (energeticp_index(:)-720)/2;
23     end
24 end
25
26 [~,vyska] = size(energetic);
27 for i = 1:vyska
28 end

```

## B OBSAH PŘÍLOŽENÉHO CD

Na přiloženém médiu naleznete naprogramované skripty v Matlabu R2013a. Dále naleznete nahrávky jednotlivých poloh zdroje, jak je popsáno ve stromě souborů. Ve stromě souborů nejsou vypsány všechny nahrané soubory, princip je stejný.

```
/ ..... kořenový adresář přiloženého CD
├── Matlab_skripty ..... naprogramované skripty v Matlabu
│   ├── Nacteni ..... adresář obsahující skripty k načtení zvukových souborů
│   ├── Metoda GCC ..... adresář se skripty GCC metody
│   ├── Metoda PHAT ..... adresář se skripty PHAT metody
│   ├── Metoda ML ..... adresář se skripty ML metody
│   └── Energetická metoda ..... adresář se skripty energetické metody
├── Nahrávky
│   ├── Nahrávky bez mikrofonního pole ..... nahrávky bez mikr. pole
│   │   ├── 01_01 ..... nahrávka č. 1; 1. pozice reproduktoru
│   │   ├── Audio 9-001.wav ..... signál na 1. mikrofonu
│   │   ├── Audio 9-002.wav ..... signál na 2. mikrofonu
│   │   ├── Audio 9-003.wav ..... signál na 3. mikrofonu
│   │   ├── Audio 9-004.wav ..... signál na 4. mikrofonu
│   │   ├── Audio 9-005.wav ..... signál na 5. mikrofonu
│   │   ├── Audio 9-006.wav ..... signál na 6. mikrofonu
│   │   ├── Audio 9-007.wav ..... signál na 7. mikrofonu
│   │   └── Audio 9-008.wav ..... signál na 8. mikrofonu
│   ├── 01_02 ..... nahrávka č. 1; 2. pozice reproduktoru
│   ├── 01_03 ..... nahrávka č. 1; 3. pozice reproduktoru
│   ├── 01_04 ..... nahrávka č. 1; 4. pozice reproduktoru
│   ├── 01_05 ..... nahrávka č. 1; 5. pozice reproduktoru
│   ├── 01_06 ..... nahrávka č. 1; 6. pozice reproduktoru
│   ├── A pole ..... nahrávky A-formát signálu
│   │   └── LB 0 deg_05.wav .signál z mikrofonu Lb v pozici 0 .3 Nahrávky
│   │       s mikrofonním polem ..... nahrávky s mikr. polem
│   ├── dvere_kriz_pole..adresář s nahrávkami zvuku dveří v křížovém poli
│   │   ├── 0_stupnu ..... adresář s nahrávkami na pozici 0 stupňů
│   │   │   └── Audio-001.wav ..... nahrávka z mikrofonu na pozici 1
│   │   └── 8_stupnu ..... adresář s nahrávkami na pozici 8 stupňů
├── Návrh pole ..... adresář s návrhy polí
├── Zpracování výsledků EXCEL
├── pdf ..... jednotlivé grafické přílohy ve formátu pdf
├── ALL_methods.m ..... všechny naprogramované metody
├── grafy.m ..... tisk grafů
├── prevodzcell.m ..... převod výsledné matice z formátu cell
├── prumerovani.m ..... průměrování výsledných matic
└── small_samples.m ..... skript k testování na malých vzorkách
```

## SEZNAM SYMBOLŮ, VELIČIN A ZKRATEK

AOA	odhad úhlu přicházejícího signálu – Angle of Arrival
DOA	směr příchozího signálu – Direction of Arrival
FDOA	frekvenční rozdílnost příchozího signálu – Frequency of Arrival
GROA	rozsah vstupních úrovní přicházejícího signálu – Gain Ratios of Arrival
TSWLS	váhování nejmenšími čtverci – Weighted Least Squares
MEM	metoda maximální entropie – Maximum entropy method
SIR	odstup signálu od interference – Signal to Interference Ratio
AAR	odezva úhlových kmitočtů – Adapted Angular Response
ARMA	model autoregresivních klouzavých průměrů – Autoregressive Moving Average Model
DAS	metoda beamformingu „Zpozdí a sečti“ – Delay and Sum
FAS	metoda beamformingu „Filtruj a sečti“ – Filter and Sum
MSNR	maximální odstup signálu od šumu – Maximum Signal to Noise Ratio
MMSE	minimální střední kvadratická chyba – Minimum Mean-Squared Error
MVDR	odezva bez zkreslení s minimální odchylkou – Minimum Variance Distortionless Response
CPA	nejbližší bod příchodu signálu – Closest Point of Approach
PHAT	transformace fáze – Phase Transform
ML	pravděpodobnostní maximum – Maximum Likelihood
GCC	vzájemná křížová korelace – Generalized Cross Correlation
LCMV	lineární omezení s minimální odchylkou – Linear Constraint Minimum Variance
FFT	rychlá Fourierova transformace – Fast Fourier Transform
DFT	diskrétní Fourierova transformace – Discrete Fourier Transform
STFT	krátkodobá Fourierova transformace – Short Time Fourier Transform

iFFT	zpětná rychlá Fourierova transformace – Inverse Fast Fourier Transform
SNR	odstup signálu od šumu – Signal-to-Noise Ratio
FEKT	Fakulta elektrotechniky a komunikačních technologií
$s(t)$	signál v čase zvukového zdroje
$\psi$	váhovací filtr
$w_j$	okno pro krátkodobou Fourierovu transformaci
$f_s$	vzorkovací kmitočet
$p_{00}$	barometrický tlak
$z_a$	jednotková akustická impedance prostředí
$t$	čas
$f$	frekvence
$d$	vzdálenost
$c$	rychlost zvuku